Acc. No.....................

ISSUE LABEL
Not later than the latest date stamped below.

INTERNATIONAL SERIES IN PHYSICS

LEE A. DuBRIDGE, Consulting Editor

THE FUNDAMENTAL PRINCIPLES

OF

QUANTUM MECHANICS

This complete edition is produced in full compliance with the government's regulations for conserving paper and other essential materials.

INTERNATIONAL SERIES IN PHYSICS

LEE A. DuBRIDGE, *Consulting Editor*

THE

FUNDAMENTAL PRINCIPLES

OF QUANTUM MECHANICS

With Elementary Applications

BY

EDWIN C. KEMBLE
Professor of Physics, Harvard University

DEDICATED TO

DAYTON CLARENCE MILLER

PREFACE

This volume was originally intended to be an expansion of a summary of the elements of quantum mechanics written some years ago for the *Reviews of Modern Physics* by the author in collaboration with Professor E. L. Hill. The point of view is essentially the same as in the summary, but as the present work has grown in my hands it has lost most of its resemblance to the initial pattern.

The method of approach was dictated by the desire to meet the needs of graduate students of physics. For this reason the argument is inductive in form and applications of the theory have been interwoven with the development of the basic mathematical structure. In order to minimize the necessity for frequent consultation of mathematical reference books, a good deal of background mathematical material is included in Chap. IV.

In reading other treatises on quantum theory I have frequently been distressed by the tendency to gloss over the numerous mathematical uncertainties and pitfalls which abound in the subject. From the standpoint of the beginner there is much to be said for this practice of minimizing the defects of the theory in order to exhibit its main outlines in a compact and attractive form. Nevertheless it has seemed to me that a book which deliberately called attention to the weak spots in the argument would be of considerable value to teachers and to students of the more mature type. The work of the mathematician von Neumann provides a masterly antidote to the lack of rigor characteristic of the average physicist, but by common consent this work is too difficult for any but the most mathematical students of this subject. Hence I have been led to try my hand at bridging the gap between the exacting technique of von Neumann and the usual less rigorous formulations of the theory. In carrying this project through I have restricted the discussion to such elementary mathematical methods as are the common property of physicists today. The reader must judge my success in avoiding the Scylla of sloppy thinking and the Charybdis of tedious complexity. Fine print, starred sections, and appendices indicate portions of the material which may well be omitted or briefly scanned on first reading.

A feature of the present volume on the physical and philosophical side is its consistent emphasis on the operational point of view and on the fundamental importance of Gibbsian assemblages of independent systems in the physical interpretation of the mathematical formalism.

A considerable collection of references indicates the author's indebtedness to the ideas of others, but the list is by no means exhaustive. I have borrowed freely from other books and am particularly indebted to those of von Neumann, Dirac, and of Born and Jordan.

It is a pleasure to thank my colleagues and former colleagues, Dr. Eugene Feenberg, Dr. W. H. Furry, Professor J. C. Slater, and Professor J. H. Van Vleck for invaluable suggestions and generous assistance. I am particularly indebted to Professor Van Vleck for reading the entire manuscript and for his constant encouragement. Dr. Montgomery H. Johnson is responsible for much of the work on the continuous spectrum in Sec. 31, while Dr. Bela Lengyel and Dr. Charles H. Fay have at various times spent long hours in checking equations and other technical assistance. The author is very grateful to the librarian of the Harvard Physics Laboratory, Mrs. Miner T. Patton, for her cheerfulness and accuracy in the repeated typing of successive editions of the manuscript.

To the Milton Fund of Harvard University I am indebted for a generous grant for technical help in preparing the manuscript for the printer.

<div align="right">EDWIN C. KEMBLE.</div>

PEACHAM, VERMONT,
August, 1937.

CONTENTS

   * An asterisk before the number of a section or subsection indicates that the section
or subsection so marked may be omitted or skimmed to advantage on first reading.

NOTATION

The number of different physical and mathematical quantities to be represented by separate symbols in this book is embarrassingly large in comparison with the available letters of the Roman and Greek alphabets. For this reason the establishment of a one-to-one correspondence of symbols and meanings has proved impracticable. The author has endeavored to keep the notation consistent within each chapter and, with a few exceptions which should not be confusing, has used only one symbol for each well-defined and recurrent meaning.

The following notes may be of use to the reader who attempts to dip into the middle of the book.

An asterisk * used as a superscript denotes the complex conjugate of the number or function in question.

Ordinarily the symbol Ψ denotes a time-dependent wave function, while ψ indicates the time-free space factor of a monochromatic or single-energy Ψ. At times ψ is also used for the instantaneous form of a general wave function.

Vectors are indicated by superior arrows.

Three-dimensional vector and scalar products are indicated by the conventional $\times$ and $\cdot$, e.g., $\vec{A} \times \vec{B}$ and $\vec{A} \cdot \vec{B}$.

The scalar products of many-dimensional complex vectors and of functions are denoted by heavy parentheses, e.g.,

$$(\vec{A}, \vec{B}) = \Sigma_k A_k B_k{}^*,$$

$$(\psi(x), \varphi(x)) = \int \psi \varphi^* dx.$$

In Chap. IV the *norm* of a function f, viz., (f,f), is indicated by Nf, while the *magnitude*, or square root, of the norm is indicated by $\|f\|$.

$\overline{\Sigma}_{\alpha'}$ denotes a mixed process of summation and integration over all eigenvalue points in α'-space. *Cf.* p. 246.

Matrices are denoted by boldface type or by a typical element enclosed in double vertical rules. Thus,
$$\mathbf{H} = \|H(m,n)\|.$$
The *first* of the two indices of the typical element of an ordinary two-dimensional matrix indicates the *row*, while the *second* denotes the *column*.

The Dirac notation for an eigenfunction of α in x'-space, viz., $(x'|\alpha')$, is introduced in Sec. 36*h*, while the Dirac notation for matrix elements, *e.g.*, $(\beta''|\gamma|\beta')$, appears in Sec. 44*d*. The Dirac symbolism is employed only at points where it is particularly convenient.

REFERENCE ABBREVIATIONS

D. P.   *Differentialgleichungen der Physik*, Riemann-Weber, Braunschweig, edition of 1927.

E. Q.   *Elementare Quantenmechanik*, M. Born and P. Jordan, Berlin, 1930.

M. G. Q.   *Mathematische Grundlagen der Quantenmechanik*, J. v. Neumann, Berlin, 1932.

M. M. P.   *Methoden der Mathematischen Physik I*, R. Courant and D. Hilbert, Berlin, 2d ed., 1931.

P. Q. M.   *The Principles of Quantum Mechanics*, P. A. M. Dirac, Oxford, 1st ed., 1930; 2d ed., 1935.

Q. M.   *Quantum Mechanics*, E. U. Condon and P. M. Morse, New York, 1929.

T. A. S.   *The Theory of Atomic Spectra*, E. U. Condon and G. H. Shortley, Oxford, 1935.

FUNDAMENTAL PRINCIPLES
OF QUANTUM MECHANICS
WITH ELEMENTARY APPLICATIONS

CHAPTER I
INTRODUCTION TO THE DUALISTIC THEORY OF MATTER; DEVELOPMENT OF SCHRÖDINGER'S WAVE EQUATION

1. HISTORICAL INTRODUCTION

The first step toward the formulation of the quantum theory was made in 1900 by Max Planck in the course of a theoretical investigation of the laws of thermal radiation.[1] His problem was to explain the distribution of energy in the continuous spectrum of a heated black body as a function of its temperature. The experimental fact is that the intensity per unit frequency interval rises from zero at very low frequencies to a maximum value whose position and magnitude depend upon the temperature—then falls again, approaching zero at very high frequencies. The drop in intensity in the high-frequency region is in violent conflict with a theoretical result previously obtained by the elder Lord Rayleigh[2] on the basis of the equipartition theorem derived from the classical statistical mechanics and of the wave theory of light. Planck attributed the discrepancy to the breakdown of the equipartition theorem when applied to high-frequency oscillations and made the brilliant suggestion that, if the vibrating matter particles which emit radiation have motions restricted to certain discrete energy values, or energy levels, there would be a departure from the laws of the classical statistical mechanics of the sort required by the experimental facts. On the basis of this hypothesis Planck was able to derive a formula for the intensity of the radiation in terms of temperature and frequency which fits the empirical data within the limits of experimental uncertainty.

The acceptance of Planck's suggestion meant a complete revolution in physics since it was incompatible with both the Newtonian mechanics and the electromagnetic theory of light. As corollaries, one may infer at once that the mechanics of the collisions between atoms are completely

---

[1] M. PLANCK, *Verh. d. deut. physik. Gesell.* **2,** 237 (1900); *Ann. d. Physik* **4,** 553 (1901).

[2] RAYLEIGH, *Phil. Mag.* **49,** 539 (1900).

nonclassical and that the light is emitted and absorbed in discrete "parcels" or "quanta." A natural inference is that radiant energy is corpuscular in character. Planck himself was unwilling to entertain so radical an hypothesis and spent much time and energy in an ultimately fruitless attempt to save the wave theory of light by a modification of his original energy-level assumption. In fact so firm was the hold of the wave theory on the minds of all physicists that it was not until 1905 that Einstein,[1] then a young man of 26, seriously revived the corpuscular conception of the nature of light.

The application of Planck's hypothesis to the problem of the structure of matter was still further delayed. Although Einstein showed in 1907 that it contained the key to the problem of the low-temperature specific heats of solids, it was not until 1913 that Bohr[2] united the Rutherford nuclear conception of the atom with the energy-level hypothesis to formulate his famous theory of the structure and spectrum of hydrogen.

The initial success of the young Dane[3] was followed by a period of rather feverish, but fruitful, activity for both experimental and theoretical physicists. New discoveries followed one another in bewildering succession, and in a short time scientific understanding of the nature of matter was immeasurably deepened. Prior to 1924, however, the theoretical developments were largely of an essentially provisional character. At that time the battle between the advocates of the wave theory of light and the proponents of the corpuscular theory had led to an unsatisfactory stalemate. Those who favored the corpuscular theory had made it abundantly clear that radiation has many of the properties

---

[1] A. EINSTEIN, *Ann. d. Physik* (4) **17**, 132 (1905); **20**, 199 (1906). Einstein's first paper on the theory of relativity was sent to the publishers less than 4 months after his first paper on the corpuscular theory of light!

[2] N. BOHR, *Phil. Mag.* **26**, 476, 857 (1913).

[3] Bohr was twenty-eight years old when he published his first paper on the theory of the hydrogen spectrum. In fact, the quantum theory has been from first to last a development by young men. Einstein, as already mentioned, wrote his initial paper on the corpuscular theory of light at the age of twenty-six. Heisenberg was twenty-four years old when he laid the foundation of the matrix mechanics. Dirac and Jordan wrote their first important papers at the ages of twenty-four and twenty-three, respectively. W. Pauli, Jr., was already a figure of importance in theoretical physics when at twenty-five he formulated the exclusion principle which bears his name. Uhlenbeck and Goudsmit were twenty-five and twenty-three years of age, respectively, when they invented the spinning electron. L. de Broglie published his first paper on the wave theory of the electron at thirty-one, while Schrödinger's most important papers on wave mechanics were written at the relatively advanced age of thirty-nine. Of course the contributions of older men, especially Sommerfeld and Born, have been exceedingly valuable, but one cannot but be impressed with the importance to science of a system of education which enables young men to finish their preliminary training and start their career of productivity while the extraordinary mental energy of youth is still in full vigor.

to be expected from their model. But no satisfactory way of accounting for the characteristic wave phenomena of interference and diffraction on the basis of a pure particle theory had been found. There was abundant evidence of the reality of the energy levels postulated by Planck and Bohr—yet it had also become clear that Bohr's makeshift combination of classical mechanics and "quantum conditions" was inadequate for the working out of an exact theory of atomic structure. Moreover, aperiodic phenomena and the problem of the interaction of atoms in the formation of molecules and solids were practically untouched.

The temporary retardation in the progress of theoretical physics brought about by the limitations inherent in the Bohr theory was finally broken by the introduction of new fundamental hypotheses by Louis de Broglie[1] and Werner Heisenberg.[2] To de Broglie we owe the suggestion that matter may share the dualistic characteristics of radiation by combining the properties of waves with those of corpuscles. To Heisenberg we owe a scheme for the exact description of atomic dynamical systems by means of a new kinematics based on Bohr's "correspondence principle." De Broglie's hypothesis in the hands of Schrödinger[3] received the definitive form now generally accepted, and Heisenberg's method was converted into a powerful matrix calculus by Born and Jordan.[4] Both suggestions, despite their extreme dissimilarity, proved to be of great value. Fused into a single theory which we call the "quantum mechanics," they correct the deficiencies of the Bohr theory as a tool for investigating the structure of matter, relate the newly discovered diffraction of electron beams to the problem of locating atomic energy levels, and go a long way toward removing the dilemma regarding the nature of radiant energy.

In the form of quantum theory now most generally accepted, the dualistic nature of radiation is treated as a fact to be described rather than explained or exorcised. In accordance with de Broglie's hypothesis, a similar dualistic nature is ascribed to matter, and thus a unification in the treatment of matter and radiation is attained. The fundamental similarities between the assumed characteristics of matter and radiation form one of the most striking features of present physical theory. Differences remain, to be sure—and we can by no stretch of the imagination identify these two modes of existence—but the analogy is far reaching enough to permit the use of observations regarding the characteristics of radiation as guides in the construction of a theory of matter. The de Broglie-Schrödinger wave mechanics is the result of a conscious

[1] L. DE BROGLIE, *Nature* **112**, 540 (1923); *Thesis*, Paris, 1924; *Ann. de Physique* (10) **3**, 22 (1925).

[2] W. HEISENBERG, *Zeits. f. Physik* **33**, 879 (1925).

[3] E. SCHRÖDINGER, *Ann. d. Physik* **79**, 361, 489 (1926).

[4] M. BORN and P. JORDAN, *Zeits. f. Physik* **34**, 858 (1925).

attempt to follow such guides and affords a relatively easy method of approach to the general theory.

2. THE DUALISTIC THEORY OF RADIATION

The importance of optical analogy in the development of quantum mechanics lies primarily in the fact that the dualistic nature of light is much more obvious than the dualistic nature of matter. In the region of long waves the wavelike characteristics of radiation are strongly predominant, while in the X-ray region the corpuscular characteristics are more obvious. As the transition from one part of the spectrum to the other is continuous, the dualism is inescapable. On the other hand, serious technical difficulties stand in the way of a direct experimental study of long wave length matter waves—so that it is small wonder that the wavelike aspect of the nature of matter was discovered at a very late date.

Let us therefore begin our study of the quantum mechanics with a preliminary examination of the properties of radiation. As previously stated, the result of the conflict between the wave theory of radiation and the corpuscular theory up to 1924 was a draw. The electromagnetic theory of Maxwell gave a simple and accurate account of interference, diffraction, and dispersion, besides making proper connection with quasi-static electromagnetic phenomena in the limiting region of very long waves. The corpuscular theory gave a simple and accurate account of the fundamental laws of the photoelectric effect and the Compton effect. It could be regarded as a logical corollary of the fundamental law of spectroscopy,

$$E' - E'' = h\nu,$$

and it seemed necessary in order to account for the abrupt changes in momentum experienced by emitting and absorbing atoms and molecules in a radiation field.[1] Neither point of view gave a satisfactory description of the whole field of optics. In the case of the Doppler effect,[2] the predictions of the two theories were identical and in agreement with experiment. In some ways the two theories supplemented each other. For example, in the case of the inverse photoelectric effect (*i.e.*, the production of the continuous X-ray spectrum) the corpuscular theory was needed to account for the sharply defined high-frequency limit to the spectrum, but the help of the wave theory was needed to account for the polarization of the radiation. Either point of view gave a qualitative explanation of the variation in hardness with direction of emission, which is, in fact, a kind of Doppler effect.

[1] A. Einstein, *Physik. Zeits.* **18**, 121 (1917).
[2] E. Schrödinger, *Physik. Zeits.* **23**, 301 (1922).

To obtain a satisfactory theory of light, one must formulate a description of its behavior which embodies the characteristics of both of these conflicting points of view. As a first step toward the formulation of such a description, we observe that a similar controversy arose long ago in connection with matter. In bulk it has properties which are conveniently described from the continuum point of view. In particular, it may be the vehicle of sound waves which act like waves in a continuous medium. If matter be molecular in structure, however, we must expect this fact to be most evident in the properties of low-density gases. Experiments made on such gases do favor the molecular hypothesis and are regarded as crucial since high-density matter must in any case act in some ways like a continuum. Similarly, the corpuscular properties of radiation, if they exist, must be most evident if the corpuscles are of great energy and few in number as in the case of low-intensity beams of hard X-rays. Precisely here the evidence for atomicity through the C. T. R. Wilson ray-track experiment and the Duane-Geiger point counter is most positive and definite. On the other hand, to get a test of one of the predictions of the wave theory one must have a record of the absorption of a quantity of light containing on the basis of the particle theory a very large number of photons. Thus one may say that interference experiments show that statistically light has the properties of waves without in any way directly disproving its granular structure. In other words, the experiments which were initially regarded as evidence against the corpuscular theory are not to the point but actually show merely that the properties of the corpuscles are different from those to be expected by analogy with classical mechanics.

We then lay down as an initial postulate the hypothesis of the atomicity of radiation. As a second postulate, we assume in accordance with experiment that the wave theory gives a correct description of the average intensity distribution in ordinary interference and diffraction experiments. Thus, in predicting or describing the results of optical experiments, we make use of both concepts with the related mathematical machinery.[1] For the detailed correlation of the energy E and momentum p with the frequency ν and wave length λ we make use of the fundamental formulas due to Einstein:

$$E = h\nu, \tag{2.1}$$

$$p = \frac{h}{\lambda}. \tag{2.2}$$

---

[1] One may, if one likes, assume that light consists of *both* waves and corpuscles, or he may say that it consists of corpuscles guided by a "ghost" electromagnetic wave field [*cf.* W. F. G. Swann, *Science* **61**, 433 (1925)]. The writer would prefer to regard both waves and corpuscles ultimately as mental aids in the description and prediction of empirical results, leaving all questions regarding their objective reality to the philosophers.

Here h denotes Planck's constant as usual. These equations mean that radiation which is monochromatic from the point of view of the wave theory (*i.e.*, gives a single sharp spectrum line when analyzed by a spectrometer) consists of photons of energy $E = h\nu$ and momentum $p = h/\lambda$. We further suppose that in the case of a plane progressive wave the direction of the momentum of the associated particles is that of the forward normal to the wave front.

It is immediately evident that these assumptions create a theory of radiation which unites the partial successes of the Maxwell theory with those of the corpuscular point of view. They do not answer all questions regarding the interaction of light and matter, but that is not to be expected without a fully developed theory of matter. Neither do they make any attempt to answer the question: "Why does light act in some respects like an assemblage of corpuscles and in other respects like a spreading-wave phenomenon?" Instead, they *describe* the dualistic behavior of light. As it is now generally recognized that description rather than explanation is the true function of physical theory, this procedure is entirely correct. Furthermore, the assumptions carry with them the tacit or explicit admission that deterministic models[1] are of little use in dealing with the radiation problem. In fact, the problem of the nature of radiation as seen by physicists prior to 1925 loses its point as soon as such deterministic models are cast aside.

Statistics and indeterminism enter the theory of light when one assumes that the distribution of discrete corpuscles in space is to be calculated by means of continuous wave functions. This can only mean that the intensity of light of frequency ν in any small volume G as computed from the wave functions is a measure of the probable number of photons of energy $h\nu$ in G. If the energy in G were measured n times, the individual measurements would necessarily show fluctuations or departures from the mean—although, according to the theory, the mean itself would approach the computed value as n becomes very large.[2]

Corresponding to this theoretical indeterminism, there is an experimental indeterminism evidenced, for example, in the haphazard development of the grains on a photographic plate. The ideal test for determinism or indeterminism would be, of course, to perform repeatedly the same experiment with the "initial conditions" exactly controlled, and then see whether or not the results are identical. As this is in practice

---

[1] We here use the term "model" in a very broad sense to describe either the classical concept of a particle or of a wave with all the tacit assumptions formerly bound up with these concepts.

[2] The necessity for indeterminism is also evident if we consider the problem of reflection. Waves can always divide themselves in a definite way at the interface of two media, but each individual photon must either be reflected as a whole or pass across the boundary.

impossible, we can only examine the results under conditions as similar as possible to see whether the deviations from the mean of the results are commensurate with the uncertainty regarding the initial conditions. In the case of photons, the best we can do is to throw a beam of plane-parallel light on a small aperture and allow the emergent beam to fall upon a photographic plate. There is then an uncertainty regarding the point at which any individual developable photographic grain will appear measured by the effective diameter of the illuminated portion of the plate. This uncertainty can be decreased to a certain limiting value by decreasing the area of the aperture. But, if the opening is too small, diffraction causes the illuminated area to increase once more and thus presents a complete barrier to an indefinite reduction of the experimental uncertainty. Thus an exact control of the "future" is impossible in such an optical experiment. Whether this is due to the fact that the initial conditions for the various photons cannot be exactly repeated, or to the fact that the initial conditions do not exactly control the future, is a futile question since incapable of experimental investigation. From a practical standpoint, the field of optics is indeterministic and must remain so, unless some new mode of experiment is discovered which permits a more exact control over the behavior of photons than any we now have. . If the union of particle-theory and wave-statistics outlined above is fundamentally correct, no such mode of experiment can exist.

3. AN ANALOGY BETWEEN GEOMETRICAL OPTICS AND CLASSICAL MECHANICS

In order to formulate a theory of matter paralleling the dualistic theory of radiation sketched in the preceding section, we must invent a suitable differential equation for the wave function. This equation may be assumed to have a form similar to that of the wave equation of optics but must be so designed as to harmonize with the Newtonian mechanics in the limiting case when diffraction effects are negligible. If it is possible to set up such an equation, there should be a certain similarity between the Newtonian mechanics and diffractionless, or geometrical, optics. The desired similarity does exist and is evident at once if one compares the principle of *least time* in optics (Fermat's principle) with the principle of *least action* in mechanics.[1]

The optical principle states that the path of a ray of light (wave front normal) from a point A to a point B is always such as to give the integral

$$\int_A^B \frac{ds}{w},$$

[1] The principle of least action *originated* in an attempt by Maupertuis to obtain for the corpuscular theory of light a theorem analogous to Fermat's principle (*cf.* E. T. Whittaker, *Analytical Dynamics*, 2d ed., p. 248, Cambridge, 1917).

an extreme value (usually a minimum) with respect to the integrals over all other conceivable paths for rays of the same color, or frequency. In this formula w denotes the local phase velocity of light and is a function of the frequency and the space coordinates, say x, y, z. As the frequency ν is treated as a constant in varying the integral, and as the local wave length λ is equal to w/ν, we may substitute λ for w in the statement of the principle. Using the notation of the calculus of variations we have

$$\delta \int_A^B \frac{ds}{\lambda(\nu,\, x,\, y,\, z)} = 0. \qquad (\nu \text{ unvaried}) \qquad (3\cdot1)$$

This means that the true path of the ray is characterized by the fact that for it the first variation of the path length measured in wave lengths is zero.

The usual elementary derivation of Fermat's principle[1] is valid only for light rays moving in homogeneous media bounded by plane surfaces of discontinuity and is based on the assumptions of geometrical optics, *viz.*, that light rays travel in straight lines in such media except at the boundaries where they are regularly reflected or refracted in accordance with the sine law

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{w_1}{w_2}.$$

The extension of the law to inhomogeneous media may be obtained by a limiting process in which the inhomogeneous medium is approximated by a heterogeneous aggregation of homogeneous volume elements, each differing slightly in index of refraction from its neighbors. Since the spreading of light waves is fully determined by the "wave equation"

$$\nabla^2 \varphi = \frac{1}{w^2(x,y,z,\nu)} \frac{\partial^2 \varphi}{\partial t^2}, \qquad (3\cdot2)$$

we may conclude that Fermat's principle is deducible under suitable restrictions from this equation. The mathematical verification of this theorem will be given at a later stage of our argument (Sec. 12, p. 46).

Let us now compare Fermat's principle in the form of Eq. (3·1) with the principle of least action. In the case of a single particle of total energy E, kinetic energy T, and mass μ, moving through a force field with potential energy $V(x,y,z)$, the latter principle requires that the value of the action integral

$$S = \int_A^B 2T\, dt = \int_A^B \sqrt{2\mu(E - V)}\, ds \qquad (3\cdot3)$$

[1] *Cf.* P. DRUDE, *Theory of Optics*, Chap. I, Art. 2, English translation by Mann and Millikan, New York, 1917.

for the natural, or mechanical, path between the two points A and B shall be an extremal as compared with its values for adjacent paths and the same value of E.   For such a particle the integrand in the second form of the integral, *viz.*, $\sqrt{2\mu(E - V)}$, is the absolute value of the instantaneous momentum which the particle would assume at x, y, z if it had the energy E.   Denoting this quantity by $p(E,x,y,z)$, we may state the principle in the form (*cf.* Appendix A)

$$\delta \int_A^B p(E,x,y,z)ds = 0. \qquad (E \text{ unvaried}) \qquad (3\cdot4)$$

It will be convenient to call $p(E,x,y,z)$ the *classical local momentum* to distinguish it from the true quantum mechanical momentum to be defined in Secs. 14 and 15 and from the local momentum of Young.[1]

Evidently Eqs. (3·1) and (3·4) present formally identical mathematical problems.   In fact, the paths of particles in the Newtonian dynamics may be identified with the "rays" of a problem in geometrical optics in which the wave length is adjusted to make the integrand of the integral of Eq. (3·1) proportional to the integrand of Eq. (3·4).   To be precise, the condition imposed on the wave problem is that

$$\frac{C}{\lambda} = \frac{C\nu}{w} = p(E,x,y,z) = \sqrt{2\mu[E - V(x,y,z)]}, \qquad (3\cdot5)$$

where C may be any function of ν, and ν in turn must be a function of E.

Comparing Eqs. (3·5) and (2·2), we observe that the required relation between the classical local momentum of the particle and the corresponding local wave length is identical, except for a possible constant factor, with the relation between momentum and wave length in the dualistic theory of light.

The analogy between the variational principles of geometrical optics and particle dynamics was seized upon by Sir William Hamilton in the early part of the nineteenth century and used as a guide in the development of dynamical and optical theory.   The Hamilton-Jacobi partial differential equation is the fruit of this development, but Hamilton himself regarded the analogy as an analogy only.   It remained for de Broglie and Schrödinger to show that it may profitably be used as a stepping stone for the development of a true wave mechanics similar in form to physical rather than geometrical optics.

If Hamilton had worked a little later, he might even have discovered Schrödinger's wave equation, for he would almost certainly have been led to it, if he had sought a wave equation for his waves.   But in his time, the wave theory of light was just beginning to be investigated, and it was not yet the fashion to describe waves as solutions of a partial

[1] L. A. Young, *Phys. Rev.* **38**, 1612 (1931), **39**, 455 (1932).

differential equation.   Hamilton contented himself, as was the custom
in optics, with investigating the positions of the wave fronts of his waves.
He found the wave length of mechanical waves as it depends on position,
the index of refraction as a function of the potential energy, and so on—
the latter being simply the carrying over to mechanics of Newton's
idea of the optical index of refraction as a function of potential.   But
he did not try to set up a wave equation.   And neither, at first, did
de Broglie, led so much later to precisely similar conceptions.   The
final step of writing out such an equation for matter waves and applying
it both to large-scale mechanics and to the mechanics of atomic systems
was left to Schrödinger.

4. WAVE PACKETS AND GROUP VELOCITY

The parallelism between the principle of least action and Fermat's
principle does not complete the reduction of the laws of Newton's
mechanics to a form similar to that of the laws of geometrical optics.
These principles deal solely with the paths of particles and rays, saying
nothing about the time required to traverse the path.   In order to extend
the discussion to take in the time, we must define the relation between
the motion of a large-scale body and the corresponding matter waves in
the limiting case where a sharply defined orbit exists.   Following optical
analogy we postulate that the intensity of the matter waves associated
with any particle at any space-time point x,y,z,t measures the probability
that the particle is in the neighborhood of x,y,z at the time t.   Then
if the particle is to have a fairly definite orbit, it must be associated with
a localized wave disturbance which moves with it over a definite path.
Such disturbances are familiar to students of physical optics or of other
types of wave motion.   They are usually called "wave packets" because
they can be analyzed into superpositions of infinite plane monochromatic
waves involving a narrow range of wave lengths and directions of wave
normal, and because this analysis is of fundamental importance in under-
standing their behavior as time goes on (cf. Chap. II, Secs. 9 to 12).
Optical wave packets can be formed from monochromatic optical wave
trains if a diaphragm is used to cut them off laterally and a shutter to
cut them off longitudinally.   In a non-dispersive medium where the
speed is independent of the wave length it is not necessary to start with
a monochromatic train.   Thus Fizeau in measuring the velocity of
light used wave packets formed by white light passing through apertures
in the rim of a revolving toothed wheel.   To use this method in a medium
with appreciable dispersion it would be essential to start with approxi-
mately monochromatic light, since the different colors have different
speeds.

The essential requirements to be imposed upon a wave disturbance in
order that it shall constitute a wave packet are that (a) it shall occupy a

small volume, (*b*) it shall travel with a definite speed, and (*c*) it shall travel in a definite direction. These requirements are to a certain extent mutually contradictory, as may be proved theoretically or demonstrated by appropriate experiment. Thus, in the case of an infinite train of plane waves the direction of motion is perfectly definite, but the disturbance is not localized at all. Partial localization can be produced by allowing the beam to impinge on an absorbing diaphragm provided with an aperture which, for definiteness, we assume to be circular and of radius *R*. This localization, however, is accompanied by diffraction effects which mar the sharpness of definition of the wave normal. If the initial beam is monochromatic and impinges normally on the diaphragm, and if, in addition, the radius *R* is large compared with the wave length, the diffraction effects are small and the beam is bounded rather sharply by the edge of the geometric shadow of the aperture. The direction of motion remains quite well defined. In the optical case this phenomenon is called the "rectilinear propagation of light." But if the localization of the beam in the plane of the diaphragm is made more complete, by a gradual reduction in *R*, the diffraction effects become more and more important until, in the limiting case where *R* is much smaller than the wave length, the emergent beam takes the form of a train of hemispherical waves with no definite direction of motion at all.

Similarly we may localize the wave train longitudinally if we reduce it to finite length in any manner. If the train is monochromatic and long enough to include many waves it will travel with a fairly definite speed even in a dispersive medium. (Since the speed of material particles varies with their energy, it is clear that the speed of matter waves must vary with their frequency as in the case of optical waves in a dispersive medium.) This speed of the head and tail of such a finite train is not, however, the same as the speed of the wave crests (*phase velocity*) unless the medium happens to be non-dispersive. Usually the individual waves either gain on the head of the train or fall back toward the tail to fade out into nothingness when they reach the boundary of the train. The speed of the group of waves as a whole is called the *group velocity*[1] to distinguish it from the phase velocity or the speed of the individual waves. However, if the train is made so short that it contains only a few waves, or perhaps only a fraction of a wave, the group does not hang together, but spreads longitudinally as it proceeds as if composed of dissimilar elements traveling at different rates of speed. Consequently this second kind of localization of a wave disturbance must not be carried

[1] *Cf.* A. SCHUSTER and J. W. NICHOLSON, *Theory of Optics*, 3d ed., p. 326, London, 1928; T. H. HAVELOCK, *The Propagation of Disturbances in dispersive Media*, Cambridge, 1914. Havelock ascribes the first discussion of group velocity to Hamilton. *Proc. Roy. Irish Acad.* **1**, 267, 341 (1839).

too far if it is not to conflict with the second requirement for a wave packet.

By a suitable compromise, then, we can devise wave disturbances appropriate to a dispersive medium which satisfy all three requirements for a wave packet. Such a disturbance will necessarily have a fairly well defined wave length and direction of motion, although the Fourier analysis would resolve it into a superposition of a continuous spectrum of infinite plane wave systems whose wave lengths and wave normals are spread out over a very narrow range grouped about the wave length and wave normal appropriate to the interior of the packet. We shall return to this Fourier analysis in the next chapter. For the present it will suffice to observe that the lack of sharpness in the definitions of the position, wave length, and direction of motion of such a packet is to be correlated with a corresponding lack of sharpness in the position, momentum, and direction of motion of the associated photon or matter corpuscle. The classical velocity and orbit of the corpuscle must then be identified with the velocity and orbit of the centroid of the wave packet. Thus classical mechanics is eventually to be regarded as a limiting case of the mechanics of matter wave packets.

It follows that if the analogy between the principle of least time and the principle of least action really means anything it must be possible to show that the orbit and orbital velocity of a large-scale "particle"[1] in the Newtonian mechanics are identical with the orbit and orbital velocity of a wave packet in a suitably defined wave problem. To make use of the analogy of the preceding article we ought to show (a) that wave packets travel along the rays of geometric optics, and (b) that the speed of the packet is the same as that of the corresponding mechanical particle.

To avoid interrupting the main argument here we shall assume proposition (a) without proof[2] for the present and will base our discussion of proposition (b) on the familiar formula for the group velocity of a finite train of waves in a homogeneous dispersive medium.[3] Denoting the average, or "interior" wave length of the group or packet by $\bar{\lambda}$ and the phase velocity by w, the usual expression for the group velocity v_g is

$$v_g = w(\bar{\lambda}) - \bar{\lambda}\left(\frac{\partial w}{\partial \lambda}\right)_{\lambda=\bar{\lambda}}.$$

---

[1] The motion of the center of mass of a system of particles follows the same laws as the motion of a single particle both classically and in the quantum mechanics (*cf.* Sec. 15, p. 64).

[2] The identity of the paths of wave packets or narrow beams of light with the rays defined by the normals to the corresponding extended wave systems is commonly assumed without proof in textbooks on physical optics. The assumption is validated in Sec. 12, Chap. II.

[3] *Cf.* footnote 1, p. 11.

Using the relation between wave length, frequency, and phase velocity we readily convert this expression into the more compact form,

$$\frac{1}{v_g} = \left[\frac{\partial}{\partial \nu}\left(\frac{1}{\lambda}\right) \right]_{\lambda = \bar{\lambda}}, \tag{4.1}$$

where the differentiation is carried out with ν, x, y, z acting as independent variables. Proofs of the group-velocity formula for matter waves are given in Secs. 9, 10, and 12.

If the energy of the particle and the wave length are related as in Eq. (3.5), and if the group velocity is identified with the speed of the particle v, Eqs. (3.5) and (4.1) give

$$\frac{1}{v} = \frac{\partial}{\partial \nu}\left(\frac{p}{C}\right).$$

As p is a function of E, x, y, z and as the spatial coordinates are independent of ν

$$\frac{1}{v} = -\frac{p}{C^2}\frac{dC}{d\nu} + \frac{1}{C}\frac{\partial p}{\partial E}\frac{dE}{d\nu}.$$

But

$$\frac{\partial p}{\partial E} = \frac{\mu}{p} = \frac{1}{v};$$

hence

$$\frac{dE}{d\nu} - C = \frac{pv}{C}\frac{dC}{d\nu}. \tag{4.2}$$

The product pv depends upon x, y, and z, but the left-hand member of the above equation is independent of these variables. Hence $dC/d\nu$ must vanish. C is a constant, and Eq. (4.2) reduces to[1]

$$\frac{dE}{d\nu} = C = p\lambda. \tag{4.3}$$

The linear relationship between energy and frequency thus derived suggests the hypothesis that the Einstein energy-frequency relation (2.1) holds for matter as well as for radiation. If we assume the validity of Eq. (2.1) for matter, Eq. (4.3) requires that Eq. (2.2) shall also apply to matter.

The relation thus obtained between the classical local momentum of a particle and the length of the associated matter waves was first suggested by de Broglie.[2] It has been confirmed experimentally by the

---

[1] The proof here given of the linear relationship between E and ν is closely related to that given by F. D. Murnaghan and K. F. Herzfeld, *Proc. Nat. Acad. Sci.* **13**, 330 (1927).

[2] *Loc. cit.*, footnote 1, page 3.

electron diffraction experiments of Davisson and Germer,[1] Thomson,[2] and Rupp.[3]   When Eq. (2·2) is applied to bodies of macroscopic dimensions, the wave lengths obtained are exceedingly small.   In the case of a golf ball weighing 47 grams and traveling with a speed as low as a millimeter in 10 sec., the wave length is 1.4×10^{-26} cm.!   This means that diffraction effects are hopelessly beyond the reach of experiment in the case of large-scale bodies.   On the other hand, the computed wave length becomes appreciable if Eq.'(2·2) is applied to atomic or molecular problems.   An oxygen molecule, for example, with a speed corresponding to the mean thermal energy of 300°K., has a wave length of approximately 1.5×10^{-8} cm., while an electron with a "10-volt" kinetic energy has a wave length of 5.3×10^{-8} cm.   As these dimensions are of the order of magnitude of atomic diameters and X-ray wave lengths, it is clear that diffraction effects must play a prominent part in atomic dynamics.

It is important to note that by introducing a vector $\vec{\sigma}$ having the magnitude $1/\lambda$ and the direction of the wave normal, we can throw Eq. (2·2) into the vector form

$$\vec{p} = h\vec{\sigma}. \qquad (4\cdot4)$$

We shall call $\vec{\sigma}$ the vector wave number.   Its components σ_x, σ_y, σ_z denote the number of waves per centimeter crossed by lines parallel to the x, y, and z axes respectively.

5. THE SCHRÖDINGER WAVE EQUATION FOR A SINGLE PARTICLE

5a. The Time-free Equation.—In view of the above results we may assume that a differential equation of the type of Eq. (3·2) is valid for matter waves, the phase velocity w being determined in accordance with Eqs. (2·1) and (2·2).   Then,

$$w = \lambda\nu = \frac{E}{p} = \frac{E}{\sqrt{2\mu[E - V(x,y,z)]}}, \qquad (5\cdot1)$$

or

$$w = \frac{\tfrac{1}{2}\mu v^2 + V}{\mu v} = \frac{v}{2} + \frac{V}{\mu v}. \qquad (5\cdot2)$$

(The relativistic formulation of the theory given in the next article yields a different expression for w.)   Denoting the wave function for matter waves by Ψ, we may then write the wave equation in the form

$$\nabla^2\Psi = \frac{2\mu[E - V(x,y,z)]}{E^2}\frac{\partial^2\Psi}{\partial t^2}. \qquad (5\cdot3)$$

[1] C. Davisson and L. H. Germer, *Phys. Rev.* **30**, 705 (1927); *Proc. Nat. Acad. Sci.* **14**, 317 (1928).

[2] G. P. Thomson, *Proc. Roy. Soc.* **A117**, 600 (1928); **A119**, 651 (1928).

[3] E. Rupp, *Ann. d. Physik* **85**, 981 (1928).

As in the optical case, the differential equation is applicable to mono-chromatic, or "mono-energetic" wave functions only.

This restriction means that all solutions of Eq. (5·3) which have physical significance are also to be solutions of the differential equation for a harmonic function of t:

$$\frac{1}{\Psi} \frac{\partial^2 \Psi}{\partial t^2} = -4\pi^2 \nu^2 = -\frac{4\pi^2 E^2}{h^2}. \tag{5·4}$$

Combining Eqs. (5·3) and (5·4), we obtain

$$\nabla^2 \Psi + \frac{8\pi^2 \mu}{h^2} [E - V(x,y,z)] \Psi = 0. \tag{5·5}$$

This is the *first form of Schrödinger's wave equation* for a single particle. Since the variable parameter E enters into the equation explicitly, Eq. (5·5) really includes a whole family of differential equations for each type of potential energy function V. We shall at times refer to (5·5) as the *time-free wave equation* to distinguish it from the equation (5·10) of Sec. 5b.

Equation (5·5) has the optical analogue

$$\nabla^2 \varphi = -\frac{4\pi^2}{\lambda^2} \varphi = -\frac{4\pi^2 n^2}{\lambda_0^2} \varphi, \tag{5·6}$$

where n denotes the index of refraction and λ_0 is the wave length in a vacuum. Thus one may say that the essential feature of the Schrödinger equation is that it makes the index of refraction for matter waves at each point of space proportional to the momentum which the particle would have at that point, or to $\sqrt{2\mu(E - V)}$. In the case of a particle moving under the influence of the earth's gravitational field, for example, the index of refraction will increase as the point under consideration approaches the earth's surface. From the standpoint of the wave mechanics the parabolic path of the orbit is to be attributed to the bending of the matter waves on account of the resulting inhomogeneity of space with respect to their propagation. It is precisely analogous to the bending of the sun's rays as they pass obliquely through the earth's inhomogeneous atmosphere.

5b. The Second (General) Schrödinger Equation.—Equation (5·5) is adequate as it stands for the investigation of the energy levels (frequency values) in a one-particle atomic problem, but in the case of the wave packets of Sec. 4, and in many other cases, non-monochromatic wave functions must be used. Hence we need a more general differential equation which does not contain the parameter E or its equivalent ν. Such a differential equation becomes a necessity when we have to do with problems in which the potential energy depends upon the time

explicitly, or in which for any reason the energy of the system is not conserved. A most important example is the perturbation of an atom by an external light wave which is the basis of the theory of dispersion. Here the assumption that we have to do with a single monochromatic wave function, or a fixed linear combination of such functions, breaks down entirely.

The obvious procedure for deriving a differential equation applicable to a wave packet, or to the sum of several monochromatic waves, is to eliminate E from Eq. (5·5) by differentiation. To do this we first write Eq. (5·5) in the form

$$\left[\nabla^2 - \frac{8\pi^2\mu V}{h^2}\right]\Psi = -\frac{8\pi^2\mu E}{h^2}\Psi.$$

Applying the operator $\left[\nabla^2 - \dfrac{8\pi^2\mu V}{h^2}\right]$ to each side of the equation and reducing the right-hand member with the aid of Eq. (5·4), we obtain

$$\left[\nabla^2 - \frac{8\pi^2\mu V}{h^2}\right]^2\Psi = -\frac{16\pi^2\mu^2}{h^2}\frac{\partial^2\Psi}{\partial t^2} \tag{5·7}$$

This equation is linear and is valid for any monochromatic solution of Eq. (5·5), independent of the value of E. Hence it is satisfied by the sum of any finite or uniformly convergent infinite series of monochromatic solutions of Eqs. (5·3) and (5·5). We may therefore assume that every physically admissible wave function is a solution of Eq. (5·7). The converse proposition is not plausible, however, since the above equation is of the fourth order and admits of solutions which are not linear combinations of solutions of the family of equations (5·5). Moreover Eq. (5·7) is difficult to generalize for use in connection with nonconservative systems.[1]

Fortunately there is a simpler equation than (5·7) which is adequate for our needs. Let us assume that the values of the wave function Ψ are complex numbers, or pairs of real numbers representable as complex numbers. Let us further assume that every admissible wave function is a linear combination of monochromatic functions of the special type

$$\Psi = \psi(x,y,z)e^{-\frac{2\pi iEt}{h}}, \tag{5·8}$$

where $i = (-1)^{1/2}$ and $\psi(x,y,z)$ is in general complex.[2] This represents

---

[1] *Cf.* E. Schrödinger, *Ann. d. Physik* (4) **81**, 109 (1926).

[2] If ψ is expressed in the form

$$\psi = A(x,y,z)e^{i\varphi(x,y,z)},$$

where A and φ are real, we can resolve Ψ into real and imaginary parts by the formula

$$\Psi = A\cos\left[\frac{2\pi Et}{h} - \varphi\right] - iA\sin\left[\frac{2\pi Et}{h} - \varphi\right].$$

a standing wave system if ψ is real and a progressive wave system if ψ has an appropriate complex form. In the case of a monochromatic wave of the type of Eq. (5·8) the exponential time factor in Eq. (5·5) may be canceled out, yielding the equation

$$\nabla^2\psi + \frac{8\pi^2\mu}{h^2}[E - V(x,y,z)]\psi = 0.$$ (5·9)

The factor $\psi(x,y,z)$ is sometimes called the amplitude or "space factor" of the wave function. When we have to do with monochromatic wave functions, a knowledge of ψ is equivalent to a knowledge of the complete function Ψ, for the two functions have the same absolute value and satisfy the same differential equation. The determination of the ψ functions for any problem involves the evaluation of the corresponding energies, so that the complete functions can be set up if desired. Hence we shall speak of ψ as a "time-free monochromatic wave function," or if no ambiguity is involved, we shall apply to it the simpler term "wave function."

We can now eliminate E from Eq. (5·5) by means of the relation

$$E\Psi = -\frac{h}{2\pi i}\cdot\frac{\partial\Psi}{\partial t},$$

thus obtaining the alternative wave equation

$$\nabla^2\Psi - \frac{8\pi^2\mu}{h^2}V\Psi + \frac{4\pi\mu i}{h}\frac{\partial\Psi}{\partial t} = 0.$$ (5·10)

This latter equation is much easier to handle than (5·7) as it is of the first order in t and the second order in the space coordinates. We shall call it *Schrödinger's second equation* for a single particle. Any linear combination of solutions of the family of Eqs. (5·5) having the form (5·8) is a solution of (5·10).

Not all solutions of Eqs. (5·9) and (5·10) are of direct physical interest and one of the important problems to be dealt with in Chap. III is that of defining a suitable class of solutions useful for physical purposes which we shall describe as *physically admissible*. We assume provisionally that all such physically admissible solutions of Eq. (5·10) can be expressed as linear combinations of solutions of Eq. (5·5) which are of the form specified by (5·8).[1]

The justification of the restriction of Eq. (5·8) is that complex waves of the type which it defines are easy to handle mathematically and adequate to the needs of our problem. Our problem is, it will be remem-

---

[1] Here an integral over a continuum of solutions and the limit of an infinite series of discrete solutions are included within the scope of the phrase "linear combination of solutions."

bered, merely that of formulating a mathematical description of a type of waves which will describe the facts of classical mechanics in limiting cases where diffraction effects are negligible. There is nothing in the situation which requires Ψ to be either a scalar or any particular kind of a vector. The wave function for sound waves (density or pressure) is a scalar, whereas the waves of the Maxwell electromagnetic theory consist of two three-dimensional real vectors $\vec{\mathcal{E}}$ and $\vec{\mathcal{H}}$ which may be united into a single three-dimensional complex vector. In neither case can one formulate a single second-order partial differential equation like (5·10) which summarizes the properties of either monochromatic or non-monochromatic waves in a dispersive medium. Hence the complex waves here introduced are mathematically simpler than either sound or light waves in a dispersive medium.

Of course our freedom to use these complex waves is dependent on the fact that while the intensity of the waves, measured by $|\Psi|^2$, has direct physical meaning (*cf.* p. 6), Ψ itself does not. It follows that the complex conjugate of any wave function Ψ would serve equally well to describe the same physical situation. *We shall indicate the conjugate of any complex number by an asterisk.* Thus

$$\Psi^* = \psi^*(x,y,z)e^{\frac{2\pi i E t}{h}}. \tag{5·11}$$

Evidently Ψ^* satisfies the differential equation

$$\nabla^2\Psi^* - \frac{8\pi^2\mu V}{h^2}\Psi^* - \frac{4\pi\mu i}{h}\frac{\partial\Psi^*}{\partial t} = 0. \tag{5·12}$$

The choice of Ψ rather than Ψ^* as *the* wave function is purely a matter of convention since the same physical results would be obtained by reversing the choice.[1] We shall make use of both functions in the development of the theory.

As regards the properties of the wave equation (5·10), we may observe here that it is formally similar to the equation for the diffusion of heat since it is of the first order in t. Owing to this circumstance the complete wave function in any particular case, if analytic in t, is determined by the special form of the wave equation and by the instantaneous form of Ψ at some initial instant, say $t = 0$.[2] On the other hand,

---

[1] The opposite convention has been adopted by many writers and was used by Prof. E. L. Hill and the author in their articles "General Principles of Quantum Mechanics," *Rev. Mod. Phys.* **1**, 157 (1929), **2**, 1 (1930). The present choice is sanctioned by convenience and more general use.

[2] To prove the above proposition we must show that if any two solutions of Eq. (5·10), say Ψ_1 and Ψ_2, have identical values at $t = 0$, their difference is identically zero. Let Ψ denote the difference function $\Psi_1 - \Psi_2$. Then Ψ is a solution of

the imaginary coefficient of $\partial\Psi/\partial t$ in Eq. (5·10) gives its solutions an undamped wave form quite different from the solutions of the differential equation for thermal diffusion.

*6. THE APPLICATION OF THE RESTRICTED RELATIVITY PRINCIPLE

Historically the formulation of the Schrödinger wave equations was antedated by de Broglie's application of the restricted relativity principle to the problem of the correlation of waves and free particles.[1] His argument will be reviewed here, since it leads to a very different expression for the phase velocity from that given in Eq. (5·2).

Symmetry demands that a stationary particle be associated with a stationary rather than a progressive wave system. Hence we may postulate the form

$$\Psi = f(x_0, y_0, z_0)e^{-2\pi i \nu_0 t} \qquad (6·1)$$

for the wave function of a particle referred to a system of coordinates x_0, y_0, z_0 with respect to which it is at rest. To get the corresponding wave form for a free particle moving in the direction of the z axis with a speed v, de Broglie applies the Lorentz transformation

$$z_0 = \frac{z - vt}{\sqrt{1 - \dfrac{v^2}{c^2}}}; \qquad t_0 = \frac{t - \dfrac{vz}{c^2}}{\sqrt{1 - \dfrac{v^2}{c^2}}}, \qquad x_0, y_0 = x, y,$$

which yields

$$\Psi = f\left(x, y, \frac{z - vt}{\sqrt{1 - v^2/c^2}}\right)e^{\left[\frac{-2\pi i \nu_0}{\sqrt{1 - v^2/c^2}}\left(t - \frac{vz}{c^2}\right)\right]} \qquad (6·2)$$

This expression may be made to describe an infinite plane wave system or a localized disturbance according to the hypothesis regarding the space factor $f(x_0, y_0, z_0)$. In either case the frequency defined by the phase factor is

$$\frac{\nu_0}{\sqrt{1 - v^2/c^2}}.$$

(5·12) which vanishes identically in x, y, z at $t = 0$. Thus

$$\left(\frac{\partial\Psi}{\partial t}\right)_{t=0} = 0.$$

Differentiating both sides of (5·10) with respect to t, inverting the order of differentiation and setting $t = 0$, we see that $(\partial^2\Psi/\partial t^2)_0$ must also vanish. In the same way we can prove that all the time derivatives of Ψ vanish at $t = 0$. Thus all terms in the expansion of Ψ in powers of t vanish for all values of x, y, z and hence Ψ must be identically equal to zero if it is analytic in t (cf. Sec. 32b).

[1] *Loc. cit.*, footnote 1, p. 3.

If the zero level of energy is fixed in accordance with the usual relativity expression

$$E_r = \mu c^2 = \frac{\mu_0 c^2}{\sqrt{1 - v^2/c^2}},\tag{6·3}$$

E_r and ν transform in the same way so that the fundamental relation (2·1) is unaffected by the change of coordinate systems. This fact was the first great contribution of de Broglie. The wave length for a free particle required by Eq. (6·2) is

$$\lambda = \frac{c^2}{v} \frac{\sqrt{1 - v^2/c^2}}{\nu_0} = \frac{c^2 h}{E_r v} = \frac{h}{\mu v},$$

in agreement with Eq. (2·2). The phase velocity on the other hand is

$$w = \lambda \nu = \left(\frac{h}{\mu v}\right)\left(\frac{\mu c^2}{h}\right) = \frac{c^2}{v}.\tag{6·4}$$

The discrepancy between this expression for the phase velocity and the value $v/2$ derived for a free particle from the nonrelativistic point of view [cf. Eq. (5·2)] suggests that absolute phase velocity is without physical significance. As a matter of fact there is no way in which it can be measured experimentally.

If the particle is not free, but moves under the influence of a force field with potential energy $V(x,y,z)$ we can still take into account the relativistic variation of mass with speed. The principle of least action in the form of Eq. (3·4) is still valid (cf. Appendix A), although the expression for the momentum in terms of the potential and total energies has the more complicated form

$$p(E_r,x,y,z) = \frac{1}{c}\sqrt{(E_r - V)^2 - \mu_0^2 c^4}.\tag{6·5}$$

A comparison of this principle with Fermat's principle shows that the mechanical orbits and the rays of the wave problem agree if $C/\lambda = p$, where C is a constant as in Eq. (3·5). As before we find that

$$\frac{\partial p}{\partial E_r} = \frac{\mu}{p} = \frac{1}{V}.\tag{6·6}$$

It follows from Eq. (4·1) that the group velocity of the waves is equal to the speed of the particle if we identify C with $\partial E_r/\partial \nu$ or h.

The monochromatic wave equation corresponding to the above expression for p is

$$\nabla^2\Psi = -\frac{4\pi^2}{h^2}p^2\Psi = -\frac{4\pi^2}{h^2 c^2}[(E_r - V)^2 - \mu_0^2 c^4]\Psi.\tag{6·7}$$

Introducing the ordinary energy $E = E_r - \mu_0 c^2$ we readily reduce (6·7) to the form[1]

$$\nabla^2 \Psi + \frac{8\pi^2 \mu_0}{c^2}(E - V)\Psi = -\frac{4\pi^2}{h^2 \mu_0 c^2}(E - V)^2 \Psi. \qquad (6\cdot8)$$

The right-hand member here appears a correction term to (5·5) which is usually small. We shall use this equation for the relativistic treatment of hydrogenic atoms in Chap. XIII. It is to be observed, however, that the assumption that the forces are derivable from a potential energy function is not valid in relativistic dynamics except in certain special cases and then only for a single frame of reference.[2] Hence a thoroughgoing relativistic wave equation cannot be formulated on the basis of the action function given by (6·5). Moreover, the fact that the energy E enters nonlinearly into (6·8) makes a difficulty which ultimately spoils the possibility of basing a satisfactory quantum mechanics on (6·8). For this reason we shall not take the space here to discuss the extension of the wave equation (6·7) to the non-monochromatic case.

7. THE WAVE EQUATION FOR A SYSTEM OF MANY PARTICLES

7a. Formulation of Equation.—Equations (5·5) and (5·10) can be generalized without difficulty to include the case of a system of n particles moving under the influence of conservative forces. The generalization, however, involves one important break with optical analogy. In optics a wave function spread out in ordinary three-dimensional space can describe the statistical behavior of any number of coexistent photons. The photons apparently exert no influence upon each other so that the form of an interference pattern is independent of the intensity of the light used. We infer that the same wave function may be used to describe the behavior of one photon or ten thousand. In the case of matter corpuscles, however, we must consider the "forces" which they exert on each other. In a system of n particles the motion of each depends on the coordinates of all,[3] since the potential energy V is a function of the coordinates of all.

Thus the wave function Ψ which is to describe the behavior of a system of n particles must depend on the $3n$ coordinates of the system and the time. We may say, if we like, that it is spread out over a $3n$-dimensional "coordinate" or "configuration" space in which each point represents a possible configuration of the system as a whole. But the use of such geometrical language is not essential and means

---

[1] *Cf.* O. KLEIN, *Zeits f. Physik* **37**, 895 (1926); W. GORDON, *Zeits. f. Physik* **40**, 117 (1926).

[2] *Cf.* W. PAULI, JR., "Relativitätstheorie," *Encyklopaedie der Mathematischen Wissenschaften*, **XIX**, p. 678.

[3] Classically, and hence also in the quantum mechanics.

merely that, since Ψ depends on $3n$ independent variables, it could be laid out as a "point" function only in a space having the corresponding number of dimensions. In practice the wave function for any particular case is always derived and applied by purely analytical methods independent of the concept of configuration space.

The generalization of Eqs. (5·5) and (5·10) for a system of n particles can be derived with the aid of the appropriate form of the principle of least action and a suitable extension of Fermat's principle to $3n$-dimensional waves.[1] Inspection of the three-dimensional equations suffices, however, to suggest the correct generalization. Let Eq. (5·5) be written first in the form

$$\left(\frac{1}{\mu}\frac{\partial^2}{\partial x^2} + \frac{1}{\mu}\frac{\partial^2}{\partial y^2} + \frac{1}{\mu}\frac{\partial^2}{\partial z^2}\right)\Psi + \frac{8\pi^2}{h^2}[E - V(x,y,z)]\Psi = 0. \qquad (7\cdot1)$$

It will be observed that each coordinate enters the equation through the potential energy function V and also through the corresponding operator $\dfrac{1}{\mu}\dfrac{\partial^2}{\partial x^2}$, $\dfrac{1}{\mu}\dfrac{\partial^2}{\partial y^2}$, or $\dfrac{1}{\mu}\dfrac{\partial^2}{\partial z^2}$, as the case may be. We assume that the classical potential energy function for the problem in hand is known and use it also in the wave equation. The obvious procedure for generalizing Eq. (7·1) is then to add a corresponding differential operator for each added coordinate. Let the ordinary three-dimensional Cartesian coordinates of the particles be labeled as a single series x_1, x_2, $\cdots$, x_{3n}, and let the corresponding masses be μ_1, μ_2, $\cdots$, μ_{3n}, where, of course, the three masses of any one particle are the same. Then the expanded equations, analogous to (5·5) and (5·10), are

$$\left(\sum_{k=1}^{3n}\frac{1}{\mu_k}\frac{\partial^2}{\partial x_k^2}\right)\Psi + \frac{8\pi^2}{h^2}[E - V(x_1 \cdots x_{3n})]\Psi = 0; \qquad (7\cdot2)$$

$$\left(\sum_{k=1}^{3n}\frac{1}{\mu_k}\frac{\partial^2}{\partial x_k^2}\right)\Psi - \frac{8\pi^2 V(x_1 \cdots x_{3n})}{h^2}\Psi + \frac{4\pi i}{h}\frac{\partial \Psi}{\partial t} = 0. \qquad (7\cdot3)$$

As noted above, these equations may be justified in the limiting case where diffraction effects are negligible by showing that they are in harmony with the principle of least action. An alternative procedure is to use Eq. (7·3) to derive Hamilton's canonical classical equations of motion for sharply defined wave packets (cf. Chap. VIII, Sec. 39c). The final, general justification of Eqs. (7·2) and (7·3) must obviously come through the agreement between results derived from them and the facts of experiment in the domain of atomic and molecular physics.

[1] Cf. E. C. KEMBLE, Rev. Mod. Phys. 1, 166 (1929).

Implicit in the above extension of the wave equations is the assumption that the total energy of the system of particles and the frequency of the associated wave system are related by the rule (2·1),

$$E = h\nu.$$

By means of this rule we can pass from Eq. (7·2) to Eq. (7·3) or from Eq. (7·3) to Eq. (7·2). Thus, starting from Eq. (7·3), let us seek a monochromatic solution of the form

$$\Psi = \psi(x_1, x_2, \cdots x_{3n})e^{-2\pi i\nu t}.$$

Substitution of this expression yields

$$\left(\sum_{k=1}^{3n}\frac{1}{\mu_k}\frac{\partial^2}{\partial x_k^2}\right)\Psi + \frac{8\pi^2}{h^2}(h\nu - V)\Psi = 0.$$

With the aid of (2·1) this reduces to Eq. (7·2).

7b. Relation of Schrödinger Equation to the Classical Hamiltonian Function.—Equation (7·2) is formally related to the classical Hamiltonian function for the system under consideration and can be deduced from it by the following rule of thumb. First set up the energy equation using the classical Hamiltonian function in Cartesian coordinates:[1]

$$H(p,q) \equiv \sum_{k=1}^{3n}\frac{p_k^2}{2\mu_k} + V(x_1 \cdots x_{3n}) = E. \tag{7·4}$$

---

[1] The condition for the existence of a classical Hamiltonian function is that the equations of motion of the system under consideration are reducible to the Lagrangian form

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) = 0, \qquad k = 1, 2, \cdots, f$$

[cf. Eqs. (A9), Appendix A] where L is a suitably defined function of the generalized coordinates $q_1, q_2, \cdots, q_f$, their velocities, and the time. If the kinetic energy T is a homogeneous quadratic function of the velocities as in the Newtonian nonrelativistic dynamics, and if the forces acting on the system are derivable from a potential function $V(q_1, \cdots, q_f, t)$ which does not depend on the velocities, such a reduction of the equations of motion is possible with

$$L \equiv T - V.$$

In more general cases, e.g., when the relativistic variation of mass is taken into account, or when the particles are acted on by forces of magnetic origin depending directly on the velocities, it is necessary to invent an appropriate Lagrangian function L.

When the equations have been thrown into Lagrangian form, the momenta conjugate to the coordinates $q_1, \cdots, q_f$ are defined by the equations

$$p_k = \frac{\partial L}{\partial \dot{q}_k}, \qquad k = 1, 2, \cdots, f$$

and the Hamiltonian function $H(p,q,t)$ is derived from the function $\Sigma_k p_k \dot{q}_k - L(q,\dot{q},t)$ by eliminating the velocities with the aid of the corresponding momenta.

In the Newtonian theory of a conservative system where $L = T - V$, the momen-

Second convert each side of Eq. (7·4) into an *operator*[1] by the substitution

$$p_k \rightarrow \frac{h}{2\pi i} \frac{\partial}{\partial x_k}, \qquad k = 1, 2, \cdots, 3n$$

$V \rightarrow$ operation of multiplying by V,

$E \rightarrow$ operation of multiplying by E.

Finally, let each member of the operator equation act on

$$\Psi(x_1, \cdots, x_{3n}, t).$$

The resulting differential equation is equivalent to (7·2) and is frequently expressed in the operator form

$$H\left(\frac{\partial}{\partial x_k}, x_k\right)\Psi = E\Psi, \tag{7·5}$$

where $H(\partial/\partial x_k, x_k)$, or $H(\partial/\partial x, x)$, stands for the operator obtained from the classical Hamiltonian by the above substitution. In case no ambiguity is involved, this operator, or any modification of it which plays the same role in the theory, is designated by the simple symbol H.

To set up Eq. (7·3) one proceeds as above, except that one replaces E by the operator $-\dfrac{h}{2\pi i} \dfrac{\partial}{\partial t}$.

*7c. The Schrödinger Equation and the Hamilton-Jacobi Equation.— These substitutions form an obvious parallel to those which are made in setting up the Hamilton-Jacobi partial differential equation of classical dynamics.[2] The latter equation has two forms similar to (7·2) and (7·3) respectively. In case the Hamiltonian function H does not depend explicitly on the time, it remains constant during any natural motion of the system and is identified with the energy of the system. The first form of the Hamilton-Jacobi equation is obtained from the energy equation $H(p,q) = E$ by means of the substitution

$$p_k \rightarrow \frac{\partial S}{\partial q_k}, \qquad k = 1, 2, \cdots, f$$

---

tum p_k reduces to $\partial T/\partial \dot{q}_k$, and the Hamiltonian function is equal to the total energy E. In the case of a system of n particles with Cartesian coordinates, p_k becomes $\mu_k \dot{x}_k$ and the Hamiltonian function reduces to the simple form (7·4).

Whatever the special forms of L and H may be, the f second-order Lagrangian equations given above are equivalent to the $2f$ first-order canonical equations of Hamilton

$$\frac{dq_k}{dt} = \frac{\partial H}{\partial p_k}; \qquad \frac{dp_k}{dt} = -\frac{\partial H}{\partial q_k}.$$

[1] An operator is a rule for the transformation of one function into another. The transformed function, or "transform," can have the same arguments as the original function, or a different set.

[2] *Cf.* any standard text in advanced analytical dynamics.

Any system of generalized coordinates q_1, q_2, $\cdots$, q_f is permissible. A solution of the resulting inhomogeneous first-order partial differential equation

$$H\left(\frac{\partial S}{\partial q_1}, \cdots, \frac{\partial S}{\partial q_f}, q_1, \cdots, q_f\right) = E \qquad (7\cdot6a)$$

is sometimes called an "action function," although it is not to be identified with the action integral of the principle of least action (*cf.* footnote 1, p. 44).

The second form of the Hamilton-Jacobi equation is applicable even when H involves t explicitly and the energy is not conserved. It is obtained from $H(p,q,t)$ by replacing p_k by $\partial A/\partial q_k$ for all values of k and equating the resulting expression to $-\partial A/\partial t$. Thus

$$H\left(\frac{\partial A}{\partial q_1}, \cdots, \frac{\partial A}{\partial q_f}, q_1, \cdots, q_f, t\right) = -\frac{\partial A}{\partial t}. \qquad (7\cdot6b)$$

The parallelism between this classical equation and the wave equation

$$H\left(\frac{h}{2\pi i}\frac{\partial}{\partial x_1}, \cdots, \frac{h}{2\pi i}\frac{\partial}{\partial x_f}, x_1, \cdots, x_f\right)\Psi = -\frac{h}{2\pi i}\frac{\partial\Psi}{\partial t} \qquad (7\cdot3a)$$

is striking. We have not followed Schrödinger[1] in deriving the general wave equation $(7\cdot3a)$ directly from $(7\cdot6b)$ but the reader will find a discussion of the intimate relation between these equations in Chap. II, Secs. 11, 12.[2] The apparently formal parallelism between the general wave equation and the second form of the Hamilton-Jacobi equation actually gives a basis for proving the asymptotic agreement between the classical mechanics and quantum mechanics in cases which lie outside the range of the principle of least action, *e.g.*, in the case of nonconservative systems where the forces depend on the velocities or are explicit functions of the time. Furthermore it resolves the apparent ambiguity in the sign of the exponent in Eq. (5·8). If this sign is negative, we must use the substitution $E \rightarrow -\dfrac{h}{2\pi i}\dfrac{\partial}{\partial t}$ in setting up (7·3) or (7·3a), whereas a positive sign would correspond to the substitution $E \rightarrow \dfrac{h}{2\pi i}\dfrac{\partial}{\partial t}$. If the Hamiltonian operator $H\left(\dfrac{h}{2\pi i}\dfrac{\partial}{\partial q}, q\right)$ is real, the two choices are equivalent as indicated by Eq. (5·12). However, in the more general case of a complex Hamiltonian operator, such as the one to be introduced in Eq. (7·8), they are not equivalent. In order to maintain the parallelism between (7·3a) and (7·6b) in such cases it is necessary to use the sub-

---

[1] *Cf.* footnote 3, p. 3.
[2] *Cf.* also G. D. BIRKHOFF, *Proc. Nat. Acad. Sci.* **19**, 339 (1933).

stitution $E \rightarrow -\dfrac{h}{2\pi i}\dfrac{\partial}{\partial t}$ along with $p_k \rightarrow \dfrac{h}{2\pi i}\dfrac{\partial}{\partial q_k}$ or else reverse signs throughout.[1]

From one point of view the sign of the operator substituted for E is somewhat puzzling on first examination. It is well known that the function S of Eq. (7·6a) is the generating function of a contact transformation which replaces the variables $q_1, \cdots, q_f$ by a new set one of which is the time, or differs from it by a constant. The momentum conjugate to t in this new set of coordinates is E. Since we substitute $+\dfrac{h}{2\pi i}\dfrac{\partial}{\partial x_k}$ for the momentum conjugate to x_k, one might argue that to be consistent we ought to substitute $+\dfrac{h}{2\pi i}\dfrac{\partial}{\partial t}$ for E. The argument is fallacious, however, for the classical Hamiltonian from which we form the Hamiltonian operator in Eq. (7·3a) has *not* been subjected to the above-mentioned transformation. E is *not* conjugate to t when the independent variables are $q_1, \cdots, q_f$ and we cannot make substitutions corresponding to two different sets of independent variables in the same equation.

It is possible, however, to get a satisfactory classical analogue for the substitutions used in setting up Eq. (7·3a) without using the Hamilton-Jacobi equation. To do so one makes use of a classical scheme for treating the time on the same formal basis as the spatial coordinates.[2] In this scheme a parameter τ is introduced as the independent variable and the number of independent coordinates for an f-dimensional system is stepped up to $f + 1$ by adding the time to the ordinary coordinates $q_1, \cdots, q_f$. $2f + 2$ Hamiltonian equations are then set up with $-E$ playing the role of momentum conjugate to the coordinate t.

*7d. The Wave Equation for a System of Charged Particles in a Classical External Electromagnetic Field.—For use in connection with the study of the Zeeman and Stark effects and for a discussion of the absorp-

---

[1] It is much easier to transform the Hamiltonian operator and the first-order Hamilton-Jacobi equation from one coordinate system to another than to make the corresponding transformation of the second-order wave equation (7·3a). Hence there is an obvious temptation to set up the wave equation in a generalized coordinate system p,q by transforming the classical Hamiltonian and subsequently making the substitution

$$p_k \rightarrow \frac{h}{2\pi i}\frac{\partial}{\partial q_k}: \qquad E \rightarrow -\frac{h}{2\pi i}\frac{\partial}{\partial t}.$$

This is not permissible in general, however, and it is necessary to set up the wave equation in Cartesian coordinates first and apply the transformation afterward, or to make use of some method definitely proved to be equivalent (*cf.* VII, Sec. 35b).

[2] See the article on Hamilton-Jacobi theory by L. Nordheim and E. Fues, in Geiger and Scheel's *Handbuch der Physik* Band V, Kap. 3, Ziff. 4, Berlin, 1927.

tion of radiation we shall need a wave equation applicable to a system of charged particles in a classical electromagnetic field. Although the derivation of this equation is somewhat technical for an introductory chapter, we insert it here to avoid repetition in later chapters.

As we have introduced the subject of wave mechanics by a consideration of the corpuscular theory of light, *i.e.*, with a preliminary study of the quantum properties of the electromagnetic field, the reader will be inclined to raise his eyebrows at the attempt to combine a quantum theory of the atom with a classical picture of an interacting electromagnetic field. Our excuse for the construction of such a hybrid theory lies partly in the extreme difficulty of formulating a satisfactory thoroughgoing quantum theory of the interaction between matter and radiation and partly in the observation that in the limiting case of very long wave lengths—static or quasi-static fields—the corpuscular properties of the electromagnetic field recede into the background while the classical properties dominate. Hence we can reasonably hope that such a classical treatment of the field will be in asymptotic agreement with experiment as the wave lengths under consideration become very large. As a matter of fact the absorption formulas which the theory yields have proved satisfactory over a very wide range of the spectrum.

Following the method sketched above, we shall begin by constructing the appropriate classical Hamiltonian function in Cartesian coordinates, converting it into an operator as before and using the operator to form an equation of the form (7·3a). If the external field varies with the time, the Hamiltonian will involve t explicitly and we have to do with a case in which the energy is not conserved.

Consider a system of n charged particles moving in an external classical electromagnetic field with the scalar potential $\Phi(x, y, z, t)$ and the vector potential $\vec{\mathfrak{a}}(x, y, z, t)$. Let $\mathfrak{a}_x{}^{(j)}$, $\mathfrak{a}_y{}^{(j)}$, $\mathfrak{a}_z{}^{(j)}$ denote the components of the vector potential at the point x_j, y_j, z_j, where the jth particle is located. Let μ_j and e_j denote respectively the mass and the algebraic value of the charge of that particle. The classical Hamiltonian function then takes the form[1]

$$H(p,x,t) = \sum_{j=1}^{n}\left(\frac{1}{2\mu_j}\right)\left[\left(p_x{}^{(j)} - \frac{e_j}{c}\mathfrak{a}_x{}^{(j)}\right)^2 + \left(p_y{}^{(j)} - \frac{e_j}{c}\mathfrak{a}_y{}^{(j)}\right)^2 \right.$$
$$\left. + \left(p_z{}^{(j)} - \frac{e_j}{c}\mathfrak{a}_z{}^{(j)}\right)^2\right] + V(x_1, \cdots, z_n) + \sum_{j=1}^{n}e_j\Phi^{(j)}. \quad (7·7)$$

Here V denotes as usual the internal potential energy of the system which,

---

[1] *Cf.* VAN VLECK, *The Theory of Electric and Magnetic Susceptibilities*, Chap. I, p. 7, Oxford, 1932; J. FRENKEL, *Lehrbuch der Elektrodynamik*, Vol. I, pp. 330–331, Berlin, 1926.

in most cases, is computed by the electrostatic method. Φ includes only the external part of the total scalar potential.   Making the substitution $p_x{}^{(j)} \rightarrow \dfrac{h}{2\pi i}\dfrac{\partial}{\partial x_j}$, etc., we obtain the desired Hamiltonian operator

$$H\left(\frac{\partial}{\partial x}, x, t\right) = \sum_{j=1}^{n}\frac{1}{2\mu_j}\left[\left(\frac{h}{2\pi i}\frac{\partial}{\partial x_j} - \frac{e_j}{c}\mathcal{C}_x{}^{(j)}\right)^2 + \cdots\right] + V + \sum_{j=1}^{n} e_j\Phi^{(j)}.$$

$$(7\cdot8)$$

Inserting this expression into (7·3a), we deduce the generalized Schrödinger equation

$$\sum_{=1}^{n}\frac{1}{2\mu_j}\left[\left(\frac{h}{2\pi i}\frac{\partial}{\partial x_j} - \frac{e_j}{c}\mathcal{C}_x{}^{(j)}\right)^2 + \left(\frac{h}{2\pi i}\frac{\partial}{\partial y_j} - \frac{e_j}{c}\mathcal{C}_y{}^{(j)}\right)^2 + \right.$$

$$\left.\left(\frac{h}{2\pi i}\frac{\partial}{\partial z_j} - \frac{e_j}{c}\mathcal{C}_z{}^{(j)}\right)^2\right]\Psi + V\Psi + \sum_{j=1}^{n} e_j\Phi^{(j)}\Psi = -\frac{h}{2\pi i}\frac{\partial\Psi}{\partial t}. \quad (7\cdot9)$$

If we make the customary assumption that $\vec{\mathcal{C}}$ and Φ conform to the special condition

$$\operatorname{div}\vec{\mathcal{C}} + \frac{1}{c}\frac{\partial\Phi}{\partial t} = 0, \qquad (7\cdot10)$$

Eq. (7·9) takes the form

$$\sum_{j=1}^{n}\frac{1}{\mu_j}\left[\nabla_j{}^2\Psi - \frac{4\pi i e_j}{hc}\vec{\mathcal{C}}^{(j)}\cdot\nabla_j\Psi + \frac{2\pi i e_j}{hc^2}\frac{\partial\Phi^{(j)}}{\partial t}\Psi\right] - \frac{8\pi^2}{h^2}V'\Psi + \frac{4\pi i}{h}\frac{\partial\Psi}{\partial t} = 0.$$

$$(7\cdot11)$$

Here

$$V' \equiv V + \sum_{j=1}^{n}\left(e_j\Phi^{(j)} + \frac{e_j{}^2}{2c^2\mu_j}|\vec{\mathcal{C}}^{(j)}|^2\right),$$

$$\vec{\mathcal{C}}^{(j)}\cdot\nabla_j\Psi \equiv \mathcal{C}_x(x_j,y_j,z_j)\frac{\partial\Psi}{\partial x_j} + \mathcal{C}_y(x_j,y_j,z_j)\frac{\partial\Psi}{\partial y_j} + \mathcal{C}_z(x_j,y_j,z_j)\frac{\partial\Psi}{\partial z_j}.$$

In the case of a system of plane waves or a pure radiation field resolvable into a superposition of plane waves, it is always possible to choose a vector potential $\vec{\mathcal{C}}$ with zero divergence and to set Φ equal to zero. Moreover, the term $\sum \dfrac{e_j{}^2}{2c^2\mu_j}|\vec{\mathcal{C}}^{(j)}|^2$ is usually so small in practice that

it can be safely neglected. Thus the wave equation for an atomic system interacting with a pure radiation field reduces to

$$\sum_{j=1}^{n} \frac{h^2}{8\pi^2\mu_j}\left[\nabla_j{}^2\Psi - \frac{4\pi i e_j}{hc}\vec{\alpha}^{(j)} \cdot \nabla_j\Psi\right] - V\Psi = \frac{h}{2\pi i}\frac{\partial\Psi}{\partial t}. \qquad (7\cdot12)$$

In consequence of the basic relations

$$\vec{\mathfrak{K}} = \text{curl }\vec{\alpha}, \qquad \vec{\varepsilon} = -\text{grad }\Phi - \frac{1}{c}\frac{\partial\vec{\alpha}}{\partial t},$$

which give the magnetic and electric vectors $\vec{\mathfrak{K}}$ and $\vec{\varepsilon}$ in terms of the potentials $\vec{\alpha}$ and Φ, the field vectors are invariant with respect to a substitution of the form

$$\vec{\alpha} \to \vec{\alpha} + \text{grad } U(x,y,z,t). \qquad \Phi \to \Phi - \frac{1}{c}\frac{\partial U}{\partial t}. \qquad (7\cdot13)$$

Hence the physical results obtained from the wave equation (7·9) should be unchanged by the same substitution. This kind of invariance is called *gauge invariance*. As a matter of fact, if we apply the transformation (7·13) to the potentials and write

$$\Psi = \Psi_0(x_1, \cdots, z_n, t) \exp\left(\frac{2\pi i}{hc}\sum_j e_j U^{(j)}\right),$$

where Ψ_0 is a solution of (7·9), we find that Ψ is a solution of the transformed equation.[1] The phase factor $\exp\left(\frac{2\pi i}{hc}\sum_j e_j U^{(j)}\right)$ does not affect the value of $|\Psi|^2$ or the vector mass current density $\vec{I}$ whose form for the three-dimensional case is worked out in Sec. 8. We conclude that any legitimate alteration of the potentials produces a modification of the phase of the wave function which does not alter its physical implications. It will be observed that if the specialized wave equation (7·11) is used, we can make the transformation (7·13) only if

$$\nabla^2 U - \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} = 0.$$

8. THE PHYSICAL INTERPRETATION OF THE WAVE FUNCTION AND THE NORMALIZATION CONDITION

8a. Probability and Quadratic Integrability.—In concluding this chapter we return to the question of the physical significance of the Ψ

---

[1] *Cf.* V. Fock, *Zeits. f. Physik* **39**, 226 (1926); W. Pauli, Jr., "Die allgemeinen Prinzipien der Wellenmechanik," *Handbuch der Physik,* **XXIV**, Part 1, 2d ed., pp. 110–111, Berlin, 1933.

The point is easily verified by noting that

$$\left[\frac{h}{2\pi i}\frac{\partial}{\partial x_j} - \frac{e_j}{c}\left(\alpha_x{}^{(j)} + \frac{\partial U^{(j)}}{\partial x_j}\right)\right]\Psi = \exp\left(\frac{2\pi i}{hc}\sum_j e_j U^{(j)}\right)\left[\frac{h}{2\pi i}\frac{\partial}{\partial x_j} - \frac{e_j}{c}\alpha_x{}^{(j)}\right]\Psi_0$$

$$\left[\frac{h}{2\pi i}\frac{\partial}{\partial t} + \sum_j e_j\left(\Phi^{(j)} - \frac{1}{c}\frac{\partial U^{(j)}}{\partial t}\right)\right]\Psi = \exp\left(\frac{2\pi i}{hc}\sum_j e_j U^{(j)}\right)\left[\frac{h}{2\pi i}\frac{\partial}{\partial t} - \sum_j e_j\Phi^{(j)}\right]\Psi_0.$$

waves. In the case of three-dimensional waves we have already laid down the postulate that the square of the absolute value of Ψ at any space-time point x,y,z,t is to measure the probability that the associated particle is in the neighborhood of the point x,y,z at the time t. Let dF denote the probability that the particle is in the volume element $dxdydz$. To make the above hypothesis more explicit we assume that

$$dF = K|\Psi|^2 dxdydz = K\Psi\Psi^* dxdydz, \qquad (8\cdot1)$$

where K is a constant or a function of the time. Integration over all space yields the probability that the particle is somewhere, which must be unity. Hence,

$$K\iiint_\infty \Psi\Psi^* dxdydz = 1.$$

If the integral $\iiint_\infty \Psi\Psi^* dxdydz$ converges to a finite value it is equal to the reciprocal of K. Ψ is then said to be *quadratically integrable*. Many solutions of the wave equation do not satisfy this condition and for them the hypothesis $(8\cdot1)$ is inapplicable since K must vanish. Solutions of the wave equation which are not quadratically integrable are often of great mathematical and physical interest, but they play a minor role in the theory as a whole.

These solutions never correspond exactly to actual experimental physical situations. Consider, for example, the case in which Ψ is not quadratically integrable because it does not vanish rapidly enough at infinity. Then an associated particle is sure to lie outside any sphere which is drawn about the origin of the coordinate system. But in practice the apparatus and objects studied in a physical experiment must lie in a bounded portion of space, so that such a function does not fully represent a practical experimental condition. Ψ may also fail of quadratic integrability because it becomes too rapidly infinite at some finite point, say P. This occurs only when the point under consideration is a center of force such as an atomic nucleus and the Ψ function would then represent a situation in which the particle is sure to lie inside any sphere, however small, which is drawn about P as a center. In other words the particle has condensed on P, and either it doesn't exist as a separate entity or we have to do with a problem in nuclear physics with which our wave equation is not designed to deal.

Quadratically integrable wave functions can usually be normalized in such a way as to eliminate the constant K of Eq. $(8\cdot1)$. If $\Psi\Psi^*$ integrates to a constant value g^2 it is only necessary to form the new wave function

$$\Psi_1 = \frac{\Psi}{g},$$

which will also be a solution of the wave equation. Then Ψ_1 represents

the same physical situation as Ψ and in addition satisfies the normalization condition

$$\iiint_\infty \Psi_1 \Psi_1{}^* dxdydz = 1. \qquad (8 \cdot 2)$$

The constancy of $\iiint_\infty \Psi\Psi^* dxdydz$ in time is an obvious corollary on the existence of this integral in the case of *monochromatic* wave functions. In the more general case it can be proved with the aid of suitable mild restrictions on the behavior of Ψ at infinity.

8b. Normalization and Mass Current Density.—Consider the integral of $\Psi\Psi^*$ over a volume G bounded by a large sphere Λ of radius R and by small spheres $S_1, S_2, \cdot \cdot \cdot$ excluding the points at which the differential equation breaks down due to the fact that V is infinite. (We assume that V is continuous except for isolated poles.) Differentiation yields

$$\frac{\partial}{\partial t}\iiint_G \Psi\Psi^* dxdydz = \iiint_G \left(\Psi\frac{\partial \Psi^*}{\partial t} + \Psi^*\frac{\partial \Psi}{\partial t}\right) dxdydz. \qquad (8 \cdot 3)$$

Let Ψ be a solution of Eq. (7·9) for the case of a single particle of charge e. Using the time derivatives of Ψ and Ψ^* taken from this equation, we deduce the relation

$$\frac{\partial}{\partial t}\iiint_G \Psi\Psi^* dxdydz = \frac{2\pi i}{h}\iiint_G (\Psi H^*\Psi^* - \Psi^* H\Psi) dxdydz$$

$$= \frac{h}{4\pi\mu i}\iiint_G \left\{\Psi\nabla^2\Psi^* - \Psi^*\nabla^2\Psi + \frac{4\pi i e}{hc}[\Psi(\vec{\alpha}\cdot\text{grad }\Psi^*)\right.$$

$$\left. + \Psi^*(\vec{\alpha}\cdot\text{grad }\Psi) + \Psi\Psi^* \text{ div }\vec{\alpha}]\right\} dxdydz. \qquad (8 \cdot 4)$$

With the aid of Gauss's divergence theorem the volume integral on the right side of the above equation can be replaced by a surface integral. Thus, if Σ denotes the aggregate of the surfaces $\Lambda, S_1, S_2, \cdot \cdot \cdot$, and $\vec{I}$ denotes the complex vector

$$\vec{I} = \frac{h}{4\pi i}\left(\Psi^* \text{ grad }\Psi - \Psi \text{ grad }\Psi^* - \frac{4\pi i e}{hc}\vec{\alpha}\Psi\Psi^*\right), \qquad (8 \cdot 5)$$

we obtain the equation of continuity $\text{div }\vec{I} = -\mu\frac{\partial}{\partial t}\Psi\Psi^*$ and the integrated relation

$$\frac{\partial}{\partial t}\iiint_G \mu\Psi\Psi^* dxdydz = -\iint_\Sigma I_n d\Sigma. \qquad (8 \cdot 6)$$

Clearly, as the integral of $\mu\Psi\Psi^*$ over any region represents the statistical

average mass in that region, $\vec{I}$ must represent a *statistical mass current density*.

Since experience shows that matter is not created at a simple center of force we must either prove that quadratically integrable Ψ functions satisfy the condition that

$$\lim_{\rho_k \to 0} \int \int_{S_k} I_n dS = 0, \qquad (\rho_k = \text{radius of } S_k) \tag{8·7}$$

or else we must impose the condition (8·7) or equivalent additional restrictions on physically admissible wave functions.

Similarly we must rule out wave functions involving an inward or outward flow of matter from infinity. This may be done by requiring that

$$\lim_{r \to \infty} [r^2 \vec{I}(x,y,z)] = 0. \tag{8·8}$$

In this case, as in the preceding one, the restriction is a mild one which is automatically fulfilled by monochromatic wave functions.

The interpretation of the vector $\vec{I}$ as a current density, while based on the interpretation of $|\Psi|^2 dxdydz$ as probability density, is nevertheless transferable to Ψ functions which are not quadratically integrable. Mathematical simplicity in the discussion of long steady streams of independent particles is frequently obtained by treating them as if they were infinitely long and therefore contained altogether an infinite number of particles. Ψ functions which are not quadratically integrable are obviously adapted to the description of such streams, which can be regarded as the limits of finite streams as the volume over which they are steady becomes infinite. In using Ψ functions for this purpose we omit all normalization and use $|\Psi|^2 dxdydz$ as the *relative* probability of the volume element $dxdydz$ and the vector $\vec{I}$ as the *relative current density*.

In the case of the problem of n particles the wave functions have been shown to be spread over a $3n$-dimensional coordinate space. Let $d\tau$ denote the volume element $dx_1 \cdots dz_n$ in this space. If the integral $\int_\infty \Psi\Psi^* d\tau$ converges, we say that Ψ is quadratically integrable. The conditions which must be imposed on Ψ in order to insure the constancy of $\int_\infty \Psi\Psi^* d\tau$ in time are similar to those which we have derived in the three-dimensional case. (These conditions will receive further consideration in Secs. 17 and 32d.) Here we assume that these conditions

are fulfilled for all physically admissible wave functions. Such wave functions can then be normalized in accordance with the condition

$$\int_{\infty} \cdots \int \Psi\Psi^* d\tau = 1. \tag{8.9}$$

It will usually be assumed that this process of normalization has actually been carried out. By an obvious extension of our hypothesis regarding the physical interpretation of three-dimensional Ψ waves we assume in this more general case that $\Psi\Psi^* d\tau$ is the probability that the system represented by Ψ has a configuration lying in the range defined by $d\tau$.

8c. A System Consisting of Two Independent Parts.—It will be instructive to test the validity of the above assumption by the consideration of a special case in which the system consists of two independent parts. Considered separately, the two parts have wave functions Ψ_1 and Ψ_2 which satisfy the Schrödinger equations

$$H_1\Psi_1 = -\frac{h}{2\pi i}\frac{\partial\Psi_1}{\partial t}, \tag{8.10}$$

$$H_2\Psi_2 = -\frac{h}{2\pi i}\frac{\partial\Psi_2}{\partial t}. \tag{8.11}$$

On the other hand, considered as a single system, they should have a wave function Ψ which satisfies the equation

$$H\Psi = -\frac{h}{2\pi i}\frac{\partial\Psi}{\partial t}. \tag{8.12}$$

Here the Hamiltonian operator H for the united system is equal to the sum of the Hamiltonian operators H_1, H_2 of the parts. Since the probability of the simultaneous occurrence of two independent events is equal to the product of the probabilities of the individual events, the assumed physical interpretation of $|\Psi|^2$ demands that

$$|\Psi|^2 d\tau = |\Psi_1|^2 d\tau_1 |\Psi_2|^2 d\tau_2.$$

The volume element $d\tau$ in the configuration space of the united system is equal to the product of corresponding volume elements $d\tau_1$ and $d\tau_2$ in the separate configuration spaces, so that $|\Psi_1|^2|\Psi_2|^2 = |\Psi|^2$. This relation is satisfied if $\Psi = \Psi_1\Psi_2$, and substitution in (8.12) shows that $\Psi_1\Psi_2$ is actually a solution of the Schrödinger equation for the united system. We conclude that the assumed physical interpretation of $|\Psi|^2$ is in satisfactory agreement with the structure of the Schrödinger equation.

Incidentally, if the functions Ψ_1 and Ψ_2 have the monochromatic form

$$\Psi_1 = \psi_1(x_1, y_1, z_1)e^{-\frac{2\pi i}{h}E_1 t}, \qquad \Psi_2 = \psi_2(x_2, y_2, z_2)e^{-\frac{2\pi i}{h}E_2 t},$$

the product function will also be monochromatic and its time factor will be

$$\rho^{-\frac{2\pi i}{h}Et} \;=\; \rho^{-\frac{2\pi i}{h}(E_1 + E_2)t}$$

Thus the total energy of the system as defined by its vibration frequency is equal to the sum of the energies of its parts. Although this energy relationship is correct classically, it is somewhat disconcerting when interpreted as a frequency relation. To say that the frequency of the system treated as a whole is the sum of the frequencies of its parts is evidently to say that the waves have no objective reality, but are mere mathematical tools for predicting the behavior of the associated particles. The same view is supported by the discussion at the end of Sec. 7, which shows that the form of Ψ in any given case depends on the particular choice of the vector potential used in setting up the wave equations.

The author is indebted to Prof. Frenkel for the suggestion that the *fundamental* reason for the use of complex Ψ functions is that this is the only way of satisfying both the rule for combining independent probabilities and the law of addition for energy. Thus, if we assume that $\Psi = \Psi_1\Psi_2$ and give Ψ_1 and Ψ_2 the real forms

$$\Psi_1 = \psi_1 \cos\left(\frac{2\pi E_1 t}{h} + \beta_1\right), \qquad \Psi_2 = \psi_2 \cos\left(\frac{2\pi E_2 t}{h} + \beta_2\right),$$

we find that Ψ is not monochromatic but involves the two frequencies $(E_1 + E_2)/h$ and $|E_1 - E_2|/h$.

CHAPTER II

WAVE PACKETS AND THE RELATION BETWEEN CLASSICAL MECHANICS AND WAVE MECHANICS

9. WAVE PACKETS AND GROUP VELOCITY IN A ONE-DIMENSIONAL HOMOGENEOUS MEDIUM

9a. The Fourier Integral Theorem.—In this chapter we shall develop the theory of wave packets in order to exhibit more fully the asymptotic agreement between classical mechanics and wave mechanics which was assumed in setting up the Schrödinger equation in Chap. I. The theory of wave packets will then be used to formulate a more complete statistical interpretation of the Ψ waves.

We begin with a consideration of the motion of free particles in one dimension. The de Broglie waves associated with such particles move with a uniform speed independent of the coordinate x, like light waves in a homogeneous medium. Since the potential energy is constant, we may set it equal to zero without loss of generality. The Schrödinger equation (5·10) then reduces to the form

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{4\pi \mu i}{h} \frac{\partial \Psi}{\partial t} = 0. \tag{9·1}$$

A particular solution of this equation in the form of a monochromatic progressive wave is obtained by assuming that Ψ has the form

$$\Psi = e^{2\pi i(\sigma x - \nu t)}. \tag{9·2}$$

This expression satisfies (9·1) provided that the constants σ and ν are related by the formula

$$\frac{\nu}{h} = \frac{\sigma^2}{2\mu}. \tag{9·3}$$

If we interpret the frequency ν and the wave number σ in accordance with Eqs. (2·1) and (2·2), we see that the above condition is equivalent to the classical energy formula

$$E = \frac{p^2}{2\mu}.$$

A more general solution of (9·1) is obtained by taking a linear combination of such particular solutions. No such combination of discrete solutions of the above type is quadratically integrable, however, and to

35

obtain a quadratically integrable wave function we have recourse to an integral over a one-parameter family of functions of the type (9·2). For this purpose we make use of the Fourier integral theorem which is best stated for our purpose in the following form.

Let $f(t)$ denote a complex function of t whose real and imaginary parts satisfy the Dirichlet condition[1] in every finite interval and which yields a convergent integral, $\int_{-\infty}^{+\infty} |f(t)| dt$. *Then the integral*

$$g(x) = \int_{-\infty}^{+\infty} f(t)e^{2\pi i x t} dt$$

exists and implies that

$$f(t) = \int_{-\infty}^{+\infty} g(x)e^{-2\pi i x t} dx.$$

$g(x)$ is called the *Fourier transform* of $f(t)$.

If $\int_{-\infty}^{+\infty} f(t)t^2 dt$ also exists, the integral representing $g(x)$ can be differentiated under the integral sign.[2] If $f(t)$ is quadratically integrable $g(x)$ is also quadratically integrable. In fact,[3]

$$\int_{-\infty}^{+\infty} |f(t)|^2 dt = \int_{-\infty}^{+\infty} |g(x)|^2 dx.$$

[1] The Dirichlet condition is said to be satisfied by a function $f(t)$ in the interval $a < t < b$, provided that the interval can be split into a finite number of partial intervals in each of which the function is continuous and monotone. For complete rigor, $f(t)$ must be so defined at points of discontinuity that

$$f(t) = \tfrac{1}{2}[f(t + 0) + f(t - 0)].$$

[2] *Cf.* S. BOCHNER, *Vorlesungen über Fouriersche Integrale*, p. 92, Satz 35, Leipzig, 1932.

[3] The theorem is known to mathematicians as Plancherel's theorem [*cf.* M. Plancherel, *Rend. di Palermo* **30**, 289 (1910); E. C. Titchmarsh, *Proc. London Math. Soc.* (2) **23**, 279 (1923)], although the relation was first derived by the elder Lord Rayleigh on the assumption that the integrals under consideration are convergent [*Phil. Mag.* (5) **27**, 466 (1889)]. It follows from the work of Plancherel that in the proof of this theorem we may dispense with the requirement that $f(t)$ is absolutely integrable if we define the Fourier transform $g(x)$ by

$$g(x) = \frac{d}{dx} \int_{-\infty}^{+\infty} \frac{f(t)}{2\pi i t} (e^{2\pi i x t} - 1) dt.$$

A more general form of the theorem can be derived by applying it to a linear combination $f_1 + \alpha f_2$ of two quadratically integrable functions f_1 and f_2 with the Fourier transforms g_1 and g_2, respectively. It follows that for all complex values of α,

$$\alpha \int_{-\infty}^{+\infty} f_2 f_1^* dt + \alpha^* \int_{-\infty}^{+\infty} f_1 f_2^* dt = \alpha \int_{-\infty}^{+\infty} g_2 g_1^* dx + \alpha^* \int_{-\infty}^{+\infty} g_1 g_2^* dx.$$

Hence,

$$\int_{-\infty}^{+\infty} f_1 f_2^* dt = \int_{-\infty}^{+\infty} g_1 g_2^* dx.$$

We accordingly define $\Psi(x,t)$ by the equation

$$\Psi(x,t) \equiv \int_{-\infty}^{+\infty} G(\sigma)e^{2\pi i(\sigma x - \nu t)}d\sigma, \tag{9.4}$$

where $G(\sigma)$ is assumed to satisfy the conditions imposed on $f(t)$. Differentiation under the integral sign shows that $\Psi(x,t)$ is a solution of (9.1) if σ and ν are related by (9.3). If we define $\Phi(\sigma,t)$ by

$$\Phi(\sigma,t) \equiv G(\sigma)e^{-2\pi i\nu t},$$

it follows from the Fourier integral theorem that

$$\Phi(\sigma,t) = \int_{-\infty}^{+\infty} \Psi(x,t)e^{-2\pi i\sigma x}dx. \tag{9.5}$$

Moreover, it follows from Plancherel's theorem[1] that Ψ can be normalized by normalizing G. Thus

$$\int_{-\infty}^{+\infty} \Psi\Psi^* dx = \int_{-\infty}^{+\infty} \Phi\Phi^* d\sigma = \int_{-\infty}^{+\infty} GG^* d\sigma = 1. \tag{9.6}$$

Conversely, if $\Psi(x,0)$ satisfies the conditions imposed on $G(\sigma)$, and if $\Psi(x,t)$ is a solution of Eq. (9.1), it follows that $\Psi(x,t)$ is expressible in the form (9.4) and that Eqs. (9.5) and (9.6) are valid.[2]

9b. Derivation of Group-velocity Formula.—Let us next consider a special case in which $G(\sigma)$ is a function having a maximum at the point $\sigma = \sigma_0$ and approaching zero monotonically as $|\sigma - \sigma_0|$ increases. We shall suppose that G is sensibly equal to zero outside of a certain small interval M containing the point $\sigma = \sigma_0$. The wave function Ψ defined by Eq. (9.4) can then be described as a "wave packet" since it is composed of a group or packet of monochromatic progressive waves all of which have approximately the same wave number σ and the same direction of motion.

In order to study the behavior of such a packet it is convenient to resolve Ψ into its real and imaginary parts Ψ_r and Ψ_i, respectively.

---

[1] See footnote 3, p. 36.

[2] If we define $F(\sigma)$ by the equation

$$F(\sigma) \equiv \int_{-\infty}^{+\infty} \Psi(x,0)e^{-2\pi ix\sigma}d\sigma,$$

it follows from the Fourier theorem that

$$\Psi(x,0) = \int_{-\infty}^{+\infty} F(\sigma)e^{2\pi ix\sigma}d\sigma.$$

Let us form the function $\chi(x,t) \equiv \int_{-\infty}^{+\infty} F(\sigma)e^{2\pi i(x\sigma - \nu t)}d\sigma$. χ is evidently a solution of the differential equation (9.1) if σ and ν are related by (9.3) and reduces to Ψ at $t = 0$. Since the solutions of (9.1) are uniquely determined by their form at some initial instant, $t = 0$ (*cf.* footnote 2, p. 18), χ is identical with Ψ and the theorem is proved.

Let $G(\sigma)$ be real and positive. If we denote the phase angle $2\pi(x\sigma - \nu t)$ by φ we have

$$\Psi_r = \int_{-\infty}^{+\infty} G(\sigma) \cos \varphi \, d\sigma, \qquad (9\cdot7)$$

$$\Psi_i = \int_{-\infty}^{+\infty} G(\sigma) \sin \varphi \, d\sigma. \qquad (9\cdot8)$$

In evaluating these integrals x and t are to be treated as fixed parameters, while ν is a function of the variable of integration σ. The functions $\sin \varphi$ and $\cos \varphi$ are oscillatory functions of σ with the variable wave length[1]
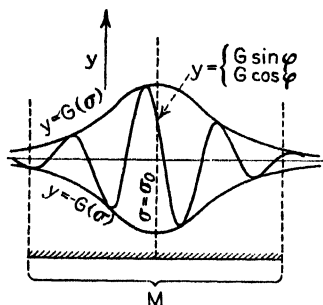
$$\lambda_\sigma = 2\pi\frac{d\sigma}{d\varphi} = \left(x - t\frac{d\nu}{d\sigma}\right)^{-1}. \qquad (9\cdot9)$$

The complete integrand of Ψ_r or Ψ_i is therefore an oscillatory function with the "envelope"

$$y = \pm G(\sigma),$$

Fig. 1.

as illustrated in Fig. 1.

If λ_σ is small in the region M, the positive and negative loops will tend to cancel each other and the integrals will vanish to a high order of approximation. In this case the monochromatic constituents of Ψ cancel each other at the space-time point x,t under consideration, "by interference," to use the language of physical optics. Actually λ_σ is small outside the neighborhood of that point σ' on the σ axis, where

$$x = t\frac{d\nu}{d\sigma}. \qquad (9\cdot10)$$

At the point in question the wave length is infinite. Then, for given values of x and t, Ψ has a maximum value if σ' is identical with σ_0, but,

[1] In the case of a monochromatic wave in a one-dimensional homogeneous medium the phase angle $\varphi(x)$ is a linear function of the coordinate x and the wave length is then defined by the relation

$$\frac{2\pi}{\lambda} = \text{coefficient of } x \text{ in } \varphi. \qquad (a)$$

It could equally well be defined by

$$\frac{2\pi}{\lambda} = \frac{\partial\varphi}{\partial x} = \text{rate of change of phase with } x. \qquad (b)$$

In the case of a nonhomogeneous medium the first definition is inapplicable and we define the local wave length by Eq. (b). Then the phase difference of two points x_1 and x_2 is given by

$$\varphi(x_2) - \varphi(x_1) = \int_{x_1}^{x_2} \frac{2\pi}{\lambda(x)} dx.$$

if σ' and σ_0 are very different, Ψ is sensibly equal to zero.   If we now allow x to vary, holding t fast, we see that we can always give it a value which will bring σ' and σ_0 together. This value is

$$\bar{x} = t\left(\frac{d\nu}{d\sigma}\right)_{\sigma_0}$$

and locates the "center" of the wave packet as a function of time. The speed with which this point moves along the x axis is called the *group velocity* and has the value[1]

$$v_g = \left(\frac{d\nu}{d\sigma}\right)_{\sigma_0} = \frac{h\sigma_0}{\mu}. \qquad (9\cdot11)$$

To determine the form of the wave packet $\Psi(x,t)$ more exactly we substitute $\sigma_0 + \eta$ for σ in (9·4), and make use of (9·3). The phase angle $\varphi = 2\pi(x\sigma - \nu t)$ becomes

$$\varphi = 2\pi\left[x(\sigma_0 + \eta) + \frac{h(\sigma_0 + \eta)^2 t}{2\mu}\right],$$

or

$$\varphi = 2\pi\left[x\sigma_0 - \frac{h\sigma_0^2 t}{2\mu} + \left(x - \frac{h\sigma_0 t}{\mu}\right)\eta - \frac{h}{2\mu}t\eta^2\right].$$

In view of (9·3) we introduce the abbreviation $\nu_0 = h\sigma_0^2/2\mu$ Then (9·4) and (9·11) yield

$$\Psi = e^{2\pi i(x\sigma_0 - \nu_0 t)}\int_{-\infty}^{+\infty} G(\sigma_0 + \eta)e^{2\pi i\left[(x - v_g t)\eta - \frac{h}{2\mu}t\eta^2\right]}d\eta. \qquad (9\cdot12)$$

However, according to hypothesis, G is negligible outside an interval M enclosing the point $\sigma = \sigma_0$. In this interval η is small. Therefore we can neglect the term $\frac{h}{2\mu}t\eta^2$ in the exponent of the integrand in (9·12), provided we restrict the discussion to values of t which are not too large. With this approximation we have

$$\Psi = e^{2\pi i(x\sigma_0 - \nu_0 t)}\int_{-\infty}^{+\infty} G(\sigma_0 + \eta)e^{2\pi i\xi\eta}d\eta;$$

where $\xi \equiv x - v_g t$. We denote the integral on the right by $u(\xi)$:

$$\Psi = u(\xi)e^{2\pi i(x\sigma_0 - \nu_0 t)}. \qquad (9\cdot13)$$

[1] Equation (9·11) is clearly equivalent to the more usual formula

$$v_g = w - \lambda\frac{dw}{d\lambda}.$$

For a more elementary discussion of group velocity see A. Schuster and J. W. Nicholson, *The Theory of Optics*, 3d ed., p. 326, London, 1928.   The writer is indebted to Prof. H. A. Kramers for suggesting the above given method of treating the subject.

$u(\xi)$ is readily seen to be a modulating amplitude factor with its maximum at $\xi = 0$, or $x = v_g t$. On either side of the center of the packet $u(\xi)$ fades off gradually toward zero. Thus the ratio of the amplitude u at an arbitrary point x to its value at the center of the packet is given by

$$\frac{u(\xi)}{u(0)} = \frac{\int_{-\infty}^{+\infty} G(\sigma_0 + \eta)e^{2\pi i \xi \eta}d\eta}{\int_{-\infty}^{+\infty} G(\sigma_0 + \eta)d\eta}.$$

Clearly this ratio will be quite small if $e^{2\pi i \xi \eta}$ goes through more than two cycles, as η ranges through the interval M. Hence we can roughly locate the "ends" of the packet by those values of ξ for which $e^{2\pi i \xi \eta}$ goes through exactly two cycles in M. Let us designate the range of values of η in M as $2\eta_1$. The "length" of the wave train is accordingly $2/\eta_1$, *i.e.*, it varies inversely with the range of wave numbers of which the train is composed.

It will be observed that the special choice of $G(\sigma)$, as real and positive, gives all the components of the packet the common phase angle zero at the space-time point $x = 0$, $t = 0$. All components can be given the arbitrary phase angle φ_0 at any point $x = x_0$, $t = t_0$ if we give G the form

$$G(\sigma) = F(\sigma)e^{i\left[\varphi_0 \ 2\pi\left(x_0\sigma \ -\frac{h\sigma^2}{2\mu}t_0\right)\right]}. \tag{9.14}$$

where F is real and positive.

It will be observed that the formula $l = 2/\eta_1$ for the length of the wave train formed by the packet depends on the special choice of the function $G(\sigma)$ and on an approximation which must break down for large values of t. When G is chosen so as to give all components the common phase angle φ_0 at $t = t_0$, the length of the wave train is a minimum at $t = t_0$. A slightly different analysis shows that the size of the train is a quadratic function of the time. For this purpose we may agree to define the length as some multiple of the root-mean-square value of $x - \bar{x}$. Denoting this root-mean-square value by Δx, we have

$$(\Delta x)^2 = \int_{-\infty}^{+\infty} (x - \bar{x})^2|\Psi|^2 dx. \tag{9.15}$$

If $G(\sigma)$ has the form given above,

$$\bar{x} = x_0 + (t - t_0)\frac{h\sigma_0}{\mu}, \tag{9.16}$$

where σ_0 is the maximum point of $F(\sigma)$. Let w denote the function

$$w = \varphi_0 - 2\pi\left(x_0\sigma - \frac{h\sigma^2 t_0}{2\mu} + \nu t\right). \tag{9.17}$$

Then

$$F(\sigma)e^{iw} = \int_{-\infty}^{+\infty} \Psi(x,t)e^{-2\pi i \sigma x}dx \tag{9.18}$$

and

$$-\frac{1}{2\pi i}\frac{\partial}{\partial\sigma}(Fe^{iw}) - \left[x_0 + (t - t_0)\frac{h\sigma_0}{\mu}\right]Fe^{iw} = \int_{-\infty}^{+\infty}(x - \bar{x})\Psi e^{-2\pi i\sigma x}dx. \quad (9\cdot19)$$

The left-hand member is readily reduced to the form

$$\left[-\frac{1}{2\pi i}\frac{dF}{d\sigma} + \frac{h}{\mu}(t - t_0)(\sigma - \sigma_0)F\right]e^{iw}.$$

Hence, by Plancherel's theorem, (9·6),

$$(\Delta x)^2 = \int_{-\infty}^{+\infty}\left[\frac{1}{4\pi^2}\left|\frac{dF}{d\sigma}\right|^2 + \frac{h^2}{\mu^2}(t - t_0)^2(\sigma - \sigma_0)^2F^2\right]d\sigma. \quad (9\cdot20)$$

This equation shows that the size of the wave packet varies quadratically with the time by an amount proportional to h^2/μ^2, and that the minimum root-mean-square

value of $x - \bar{x}$ is $\dfrac{1}{2\pi}\left[\int_{-\infty}^{+\infty}\left(\dfrac{dF}{d\sigma}\right)^2 d\sigma\right]^{\frac{1}{2}}$. Giving $F(\sigma)$ the form of a normalized Gaussian

error function, such as

$$F(\sigma) = \pi^{-\frac{1}{4}}\alpha^{\frac{1}{2}}e^{-\frac{\alpha^2(\sigma - \sigma_0)^2}{2}}, \quad (9\cdot21)$$

we find that, at $t = t_0$,

$$\Delta x = \frac{1}{\sqrt{2}}\frac{\alpha}{2\pi}. \quad (9\cdot22)$$

Our primary conclusion is that the wave packet defined by Eq. (9·4) and by the assumed form of the amplitude function $G(\sigma)$ represents a train of waves with the approximate wave length $1/\sigma_0$ [*cf.* Eq. (9·13)] moving forward individually with the phase velocity ν_0/σ_0, but having a variable amplitude such that the whole disturbance is confined to a short interval of the x axis whose center $\bar{x}$ moves forward with the group velocity v_g. Conversely, it may be proved that if at any time t, $\Psi(x,t)$ has the form just described, it must be possible to resolve it into a narrow continuous spectrum of monochromatic waves such as that defined by Eq. (9.4).

10. WAVE PACKETS IN THREE DIMENSIONS

The extension of this discussion to the motion of a free particle in *three* dimensions is very simple. The wave equation (5·10) for zero potential is solved by setting

$$\Psi = e^{2\pi i(x\sigma_x + y\sigma_y + z\sigma_z - \nu t)} \quad (10\cdot1)$$

with

$$\sigma_x{}^2 + \sigma_y{}^2 + \sigma_z{}^2 = \frac{2\mu\nu}{h}. \quad (10\cdot2)$$

This particular solution represents an infinite plane wave having the

wave-number components σ_x, σ_y, σ_z. A wave packet is obtained by forming the integral

$$\Psi(x,y,z,t) = \int\int\int_{\infty} G(\sigma_x,\sigma_y,\sigma_z)e^{2\pi i(x\sigma_x+y\sigma_y+z\sigma_z-\nu t)}d\sigma_x d\sigma_y d\sigma_z, \qquad (10\cdot3)$$

in which, for simplicity, we may assume G to be a real function which vanishes everywhere except in a small region around the point $\sigma_{0x}, \sigma_{0y}, \sigma_{0z}$ in σ space. Introducing suitable assumptions regarding the convergence of integrals of $|G|$ and $|G|^2$, when extended over all σ space, we may prove the quadratic integrability of Ψ. As before, we introduce a phase function φ defined in this case by

$$\varphi = 2\pi(x\sigma_x + y\sigma_y + z\sigma_z - \nu t) \qquad (10\cdot4)$$

and note that the integral (10·3) will be sensibly equal to zero unless x, y, z, t are given values which make the wave length λ_σ of $e^{i\varphi}$ regarded as a function of σ_x, σ_y, σ_z large at the point $\sigma_{0x}, \sigma_{0y}, \sigma_{0z}$, where G is a maximum.

It follows that the center of the packet $\bar{x}, \bar{y}, \bar{z}$ is at the point where

$$\frac{1}{\lambda_\sigma^2} \equiv \frac{1}{4\pi^2}\left[\left(\frac{\partial\varphi}{\partial\sigma_x}\right)^2 + \left(\frac{\partial\varphi}{\partial\sigma_y}\right)^2 + \left(\frac{\partial\varphi}{\partial\sigma_z}\right)^2\right] = 0 \qquad (10\cdot5)$$

for $\sigma_x = \sigma_{0x}$; $\sigma_y = \sigma_{0y}$; $\sigma_z = \sigma_{0z}$ [*cf.* Eq. (9·9)]. This is equivalent to the requirement that

$$\bar{x} - t\left(\frac{\partial\nu}{\partial\sigma_x}\right)_{\sigma_0} = \bar{y} - t\left(\frac{\partial\nu}{\partial\sigma_y}\right)_{\sigma_0} = \bar{z} - t\left(\frac{\partial\nu}{\partial\sigma_z}\right)_{\sigma_0} = 0. \qquad (10\cdot6)$$

We conclude that the center of the packet moves with a constant vector velocity whose components are

$$\left.\begin{aligned}
(v_g)_x &= \left(\frac{\partial\nu}{\partial\sigma_x}\right)_{\sigma_0} = \frac{h\sigma_{0x}}{\mu}, \\
(v_g)_y &= \left(\frac{\partial\nu}{\partial\sigma_y}\right)_{\sigma_0} = \frac{h\sigma_{0y}}{\mu}, \\
(v_g)_z &= \left(\frac{\partial\nu}{\partial\sigma_z}\right)_{\sigma_0} = \frac{h\sigma_{0z}}{\mu},
\end{aligned}\right\} \qquad (10\cdot7)$$

In other words it moves along the normal to the median wave of the packet $(\sigma_{0x}, \sigma_{0y}, \sigma_{0z})$ with the speed

$$|v_g| = [(v_g)_x{}^2 + (v_g)_y{}^2 + (v_g)_z{}^2]^{1/2} = \left(\frac{\partial\nu}{\partial|\sigma|}\right)_\sigma.$$

Because wave packets are special cases of quadratically integrable non-monochromatic solutions of the second Schrödinger equation, it

is convenient to apply the term *normal packet function* to every such solution.

*11. WAVE SURFACES AND THE HAMILTON-JACOBI EQUATION OF THE CLASSICAL DYNAMICS[1]

The theory of wave packets developed in Secs. 9–10 depends on the fact that in the case of free particles we can write down immediately an infinite family of monochromatic progressive wave solutions of the differential equation in terms of which any quadratically integrable wave function can be expressed. In case there is a variable potential energy function V, no such family of monochromatic special solutions is available. However, we can work out a family of approximate solutions which are applicable in a limited region in the neighborhood of the wave packet. With the aid of these solutions the instantaneous velocity of the packet can be determined.

As the phase velocity varies from point to point with the potential energy, the problem of setting up an approximate description of a monochromatic extended wave parallels the corresponding problem in the optics of inhomogeneous media. The approximation which we shall make is precisely the approximation involved in using the theory of geometrical optics instead of the theory of physical optics. In view of the analogy between geometrical optics and classical mechanics drawn in Sec. 3 the reader will not be surprised to discover that the same approximation leads to a connection between the wave equation and the Hamilton-Jacobi partial differential equation of the classical dynamics.

The Schrödinger equation for a single particle in three dimensions is

$$H\Psi = -\frac{h^2}{8\pi^2\mu}\nabla^2\Psi + V(x,y,z)\Psi = -\frac{h}{2\pi i}\frac{\partial\Psi}{\partial t}. \qquad (11\cdot1)$$

The desired approximate solution is obtained by assuming[2]

$$\Psi = Fe^{\frac{2\pi iA}{h}}, \qquad (11\cdot2)$$

where the "phase" A/h is a real function of x, y, z, t, while F is either real or complex but independent of the time.

Differentiation yields

$$\frac{\partial^2\Psi}{\partial x^2} = \left[-\frac{4\pi^2}{h^2}\left(\frac{\partial A}{\partial x}\right)^2 F + \frac{\partial^2 F}{\partial x^2} + \frac{4\pi i}{h}\frac{\partial A}{\partial x}\frac{\partial F}{\partial x} + \frac{2\pi i}{h}\frac{\partial^2 A}{\partial x^2}F\right]e^{\frac{2\pi iA}{h}}.$$

---

[1] The author is indebted to Prof. J. C. Slater for much of the material in this section.

[2] *Cf.* L. Brillouin, *Comptes Rendus* **183**, 24 (1926), *J. de Physique* **7**, 353 (1926); G. Wentzel, *Zeits. f. Physik* **38**, 518 (1926). The application of the substitution (11·2) to the location of energy levels is discussed in Sec. 21.

Hence Eq. (11·1) becomes

$$\left\{\frac{1}{2\mu}\left[\left(\frac{\partial A}{\partial x}\right)^2 + \left(\frac{\partial A}{\partial y}\right)^2 + \left(\frac{\partial A}{\partial z}\right)^2\right] + V + \frac{\partial A}{\partial t}\right\}F$$

$$- \frac{hi}{4\pi\mu}\left[F\nabla^2 A + 2\left(\frac{\partial A}{\partial x}\frac{\partial F}{\partial x} + \frac{\partial A}{\partial y}\frac{\partial F}{\partial y} + \frac{\partial A}{\partial z}\frac{\partial F}{\partial z}\right)\right] - \frac{h^2}{8\pi^2\mu}\nabla^2 F = 0. \quad (11·3)$$

If we think of h as a variable and allow it to approach zero we see that, in the limit, terms in h and h^2 of Eq. (11·3) drop out. The amplitude F is then arbitrary, while A must be a solution of

$$\frac{1}{2\mu}\left[\left(\frac{\partial A}{\partial x}\right)^2 + \left(\frac{\partial A}{\partial y}\right)^2 + \left(\frac{\partial A}{\partial z}\right)^2\right] + V + \frac{\partial A}{\partial t} = 0. \quad (11·4)$$

Reducing h to zero is equivalent to reducing the wave length to zero. Hence the wave function obtained by neglecting the terms of Eq. (11·3) in h and h^2 should be appropriate to the discussion of problems governed by the classical mechanics. Equation (11·4) is in fact one form of the Hamilton-Jacobi equation. The other and more familiar form is obtained by introducing the additional assumption that Ψ is monochromatic so that

$$A(x,y,z,t) = S'(x,y,z) - h\nu t = S' - Et.$$

We then have

$$\frac{1}{2\mu}\left[\left(\frac{\partial S'}{\partial x}\right)^2 + \left(\frac{\partial S'}{\partial y}\right)^2 + \left(\frac{\partial S'}{\partial z}\right)^2\right] + V = E, \quad (11·5)$$

or

$$|\operatorname{grad} S'| = \sqrt{2\mu(E - V)}. \quad (11·6)$$

In the classical mechanics a function S' which satisfies this equation is sometimes called an *action function* although it is not identical with the action function used in the formulation of the principle of least action.[1]

On the other hand, as $\frac{2\pi}{h}(S' - Et)$ is, by Eq. (11·2), the phase angle of the Ψ waves, it is evident that the "level" surfaces of S' are wave fronts. Hence it is to be expected that each such surface will generate

---

[1] Solutions of Eq. (11·5) are definite functions of x, y, z, whereas the action integral

$$S = \int_A^B 2T\,dt$$

depends upon the path of integration and is not a definite function of the coordinates of the terminal point B. If, however, we integrate along any one of the lines of the vector grad S', it may be proved that

$$S'(B) - S'(A) = \int_A^B 2T\,dt = S.$$

others in accordance with Huygens' principle. This surmise is readily shown to be correct. The vector grad S' is by definition orthogonal to the level surfaces of S' and, by Eq. (11·6), has the value $\sqrt{2\mu}(E - \overline{V})$. Hence if we choose any smooth surface as a surface of constant S', say $S' = \alpha$, we can construct a neighboring surface $S' = \alpha + d\alpha$ by going out along each normal a distance $d\alpha/|\text{grad } S'|$ and connecting the resultant terminal points (*cf.* Fig. 2); or we can make the equivalent construction by drawing spheres of radius $d\alpha/|\text{grad } S'|$ and forming their envelope. This process can be continued, setting up an infinite number of successive surfaces $S' = $ constant and so integrating the differential equation (11·5). Since the phase angle of our approximate wave function is $\dfrac{2\pi}{h}(S' - Et)$, the vector wave number is given by[1]

Fig. 2.

$$\vec{\sigma} = \frac{1}{h} \text{ grad } S', \qquad (11\cdot7)$$
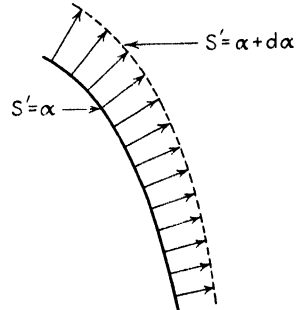
and the phase velocity is $\nu/|\vec{\sigma}|$ or $E/|\text{grad } S'|$. Hence the radii of the spheres as defined above are directly proportional to the phase velocities or wave lengths, as is proper in the Huygens construction.

In case we do not throw away the terms in h and h^2 of Eq. (11·3) we may still obtain, under certain circumstances, an excellent approximate solution by means of the substitution $\Psi = Fe^{2\pi iA/h}$ if we give F a constant value. Then all the terms in Eq. (11·3) cancel out or reduce to zero except the one involving $\nabla^2 A$. If the gradient of the potential energy $V(x,y,z)$ is small, grad S' or grad A will be nearly constant in magnitude. If the wave fronts are relatively flat, it will be nearly constant in direction. Then $\nabla^2 A = \text{div grad } S'$ will be small and the approximation good.[2]

---

[1] *Cf.* definition of wave length in footnote, p. 38.

[2] The quality of the approximation may be tested by determining the percentage discrepancy between the actual value of $\nabla^2 \Psi$ and the value required by the exact differential equation. If the amplitude F is constant, this fractional error in $\nabla^2 \Psi$ reduces to

$$\frac{h\nabla^2 A}{4\pi i\mu(E - V)}.$$

Let l, m, n denote the direction cosines of the wave normal and let T denote the kinetic energy $E - V$. Let the direction of the wave normal at the point x,y,z be that of the z axis. Then the above fraction becomes

$$\frac{1}{2\pi i}\left[\frac{\lambda}{T}\frac{\partial T}{\partial z} + \lambda\left(\frac{\partial l}{\partial x} + \frac{\partial m}{\partial y}\right) \right].$$

Thus the approximation is good provided that the fractional change in kinetic energy

A still better approximation is obtained if the imaginary terms of (11·3) are canceled out by the requirement that F satisfy the equation

$$F^2\nabla^2 A + 2F \operatorname{grad} F \cdot \operatorname{grad} A \equiv \operatorname{div}(F^2 \operatorname{grad} A) = 0. \qquad (11·8)$$

In one dimension this equation is solved by setting F equal to a multiple of $(|\operatorname{grad} A|)^{-\frac{1}{2}}$. This approximate expression for the variation in the amplitude is used in the B. W. K. method, described in Sec. 21. For our present purpose it suffices to assume that F is constant.

*12. WAVE PACKETS AND THE MOTION OF PARTICLES IN A FORCE FIELD: FERMAT'S PRINCIPLE

In Chap. I we showed that the rays determined by the principle of least time and the orbits of the classical mechanics become identical in form if the local wave length in the wave problem is related to the classical local momentum and the energy in the mechanical problem by the formula

$$\lambda = \frac{h}{\sqrt{2\mu(E - V)}}.$$

In Secs. 9 and 10 we went on to demonstrate, with the aid of certain plausible assumptions, that the motion in time of particles in a problem in classical mechanics is identical with the motion of appropriately defined wave packets in the corresponding wave problem, thus establishing an asymptotic agreement between classical mechanics and wave mechanics. The assumption that $E = h\nu$, combined with the above expression for λ, serves to define the differential equation for the wave problem (Schrödinger's equation). The remaining assumptions are as follows:

a. The rays of geometrical optics defined in terms of wave normals can be determined by the principle of least time.

b. Wave packets travel along the normals to corresponding extended waves.

c. The velocity of the wave packets is given by the group-velocity formula (4·1).

In the case of free particles the wave equation is that for a homogeneous medium and the assumptions **a** and **b** are trivial while the work of Sec. 10 verifies assumption **c**. We are now ready to check these three

---

in one wave length and the change in the direction of the wave normal in the same distance are small compared with unity. In the case of a heavy particle to which the classical mechanics should be applicable, this condition will practically always obtain owing to the small wave lengths involved. Exceptions occur when the particle is sensibly at rest and when the waves are diverging from, or converging toward, a focus.

assumptions for the general case of waves correlated with the motion of a single particle in a force field.

Let us begin with assumption **a**. We have to show that in the case of an extended monochromatic progressive wave system whose wave length is short compared with the inhomogeneity of the medium and the curvature of the wave front, lines drawn so as to be everywhere perpendicular to the wave front satisfy the equation

$$\delta \int_A^B \frac{ds}{\lambda} = 0.$$

ds denotes an element of path as usual. We may identify the wave fronts with the level surfaces of the function S' of the preceding article, and the direction of the wave normal with the direction of grad S'. Let the full line AFB connecting the wave fronts

Fig. 3.

$S' = S_A'$ and $S' = S_B'$ in Fig. 3 have the direction of grad S' at every point. For this path of integration
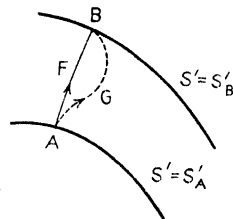
$$S_B' - S_A' = \int_{AFB} |\text{grad } S'| ds = h \int_{AFB} \frac{ds}{\lambda}. \tag{12·1}$$

Now consider any other path such as AGB. Let θ denote the angle between the path element and the vector grad S'. Then

$$S_B' - S_A' = \int_{AGB} |\text{grad } S'| \cos \theta \, ds = h \int_{AGB} \frac{\cos \theta}{\lambda} ds < h \int_{AGB} \frac{ds}{\lambda}. \tag{12·2}$$

Comparing (12·2) with (12·1), we see that $\int_A^B ds/\lambda$ has a minimum for the path of integration AFB which is everywhere directed along the wave normal. This justifies assumption **a**.

In order to validate assumptions **b** and **c** we shall extend the discussion of Sec. 10 to the case of motion in a force field by building up a wave packet from approximate wave functions of the form

$$u = \exp\left[\frac{2\pi i(S' - Et)}{h}\right], \tag{12·3}$$

where S' is a suitably chosen solution of (11·5). We know from the preceding section that there exists such a function $S'(x,y,z)$ which takes on any desired value, S_0', on any desired smooth surface Σ. Furthermore we can choose the positive sense of the gradient of S' at Σ at pleasure and can give $|\text{grad } S'|$ any value we please at any given point of space by proper choice of the energy E. Let P be an arbitrary point with the coordinates x_0, y_0, z_0. It is evident that we can choose arbitrarily S'_P, $(\text{grad } S')_P$, and the form of a smooth level surface Σ passing through P and normal to grad S' at that point. If we agree that Σ shall be

a plane and set S_P' equal to zero, there remain the three components of (grad S')$_P$, say $h\xi$, $h\eta$, $h\zeta$, at our disposal before the function S' is finally fixed. We use the three-parameter family of S' functions obtained in this way to build up the "wave packet"

$$\Psi_1 = \int \int \int_\infty G(\xi,\eta,\zeta) \exp \left\{\frac{2\pi i}{h}[S'(\xi,\eta,\zeta; x,y,z) - Et]\right\} d\xi d\eta d\zeta. \quad (12\cdot4)$$

The integrand of the above function is a solution of (11·5) and (11·4) if E is given the value

$$E = V(x_0, y_0, z_0) + \frac{h^2}{2\mu}(\xi^2 + \eta^2 + \zeta^2). \quad (12\cdot5)$$

For each set of values of ξ, η, ζ, the integrand represents a progressive wave which is plane in the neighborhood of P and has the wave number components $\sigma_x = \xi$; $\sigma_y = \eta$; $\sigma_z = \zeta$ at that point [*cf.* Eq. (11·7)].

Let $G(\xi,\eta,\zeta)$ be a real and positive function of its arguments which has a maximum at $\xi = \xi_0$, $\eta = \eta_0$, $\zeta = \zeta_0$, and vanishes outside a small region adjacent to this point. Since S' vanishes at P, the constituent waves which make up Ψ_1 are in perfect phase agreement at P, when t is zero. At this space-time point, $|\Psi_1|$ takes on its maximum value. From the argument of Sec. 9 we infer that $|\Psi_1|$ will be small at $t = 0$ and at later times except in a region K at whose center lies the point of best phase agreement.

As Ψ_1 is composed of approximate solutions of the Schrödinger equation (11·1), it must be an approximate solution itself. Consequently we can identify the velocity of the point of best phase agreement at the time $t = 0$ with the velocity of the center of the wave packet formed by the corresponding exact solution of (11·1). (In view of the alternate discussion of the motion of wave packets given in Sec. 13 it is hardly worth while to attempt any high degree of rigor in the present argument.) In order to locate the point of best phase agreement at a slightly later time Δt, we form the derivatives of the phase function $S' - E\Delta t$ with respect to ξ, η, ζ at the "point" ξ_0, η_0, ζ_0 and set them equal to zero. In calculating these derivatives we make use of the fact that S' has been so constructed that it has the expansion

$$S' = h\xi(x - x_0) + h\eta(y - y_0) + h\zeta(z - z_0) + \text{terms in higher powers} \atop \text{of } (x - x_0), \text{ etc.}$$

in the neighborhood of the point P. As we are concerned with the motion of the point of best phase agreement in a short interval of time we have to do with small displacements and can neglect the terms of higher order in $(x - x_0)$, etc. Then

$$h(x - x_0) - \Delta t\left(\frac{\partial E}{\partial \xi}\right)_{\xi_0} = 0; \qquad h(y - y_0) - \Delta t\left(\frac{\partial E}{\partial \eta}\right)_{\eta_0} = 0;$$

$$h(z - z_0) - \Delta t\left(\frac{\partial E}{\partial \zeta}\right)_{\zeta_0} = 0.$$

The components of the packet velocity are accordingly:

$$\left.\begin{aligned}
(v_g)_x &= \frac{x - x_0}{\Delta t} = \frac{1}{h}\left(\frac{\partial E}{\partial \xi}\right)_{\xi_0} = \frac{h\xi_0}{\mu}, \\
(v_g)_y &= \frac{y - y_0}{\Delta t} = \frac{1}{h}\left(\frac{\partial E}{\partial \eta}\right)_{\eta_0} = \frac{h\eta_0}{\mu}, \\
(v_g)_z &= \frac{z - z_0}{\Delta t} = \frac{1}{h}\left(\frac{\partial E}{\partial \zeta}\right)_{\zeta_0} = \frac{h\zeta_0}{\mu}.
\end{aligned}\right\} \qquad (12\cdot6)$$

These equations are identical with Eqs. (10·8) for a free particle and show that the packet moves along the normal to the "median" wave front with a speed given by the standard group-velocity formula for homogeneous media. This completes the justification of assumptions **b** and **c**.

13. DIRECT RIGOROUS PROOF OF NEWTON'S SECOND LAW OF MOTION FOR WAVE PACKETS

The work of Secs. 3, 4, 9, 10, 11, and 12 shows that localized solutions of the Schrödinger equation exist which follow the laws of ordinary mechanics. The method of developing the subject which we have used, although somewhat tedious, has the advantage of being closely related to optical theory and to the historical development of wave mechanics. It forms a natural background for the further elaboration of the quantum theory. There is, however, a more exact and elegant method of deriving the classical equations of motion for wave packets. This procedure is due primarily to Ehrenfest.[1] The method is applicable to the case of a system of particles, but for simplicity the present discussion is restricted to the case of a single particle. In Sec. 39c, the more general problem will be dealt with by a still different method.

Let Ψ denote a quadratically integrable and normalized solution of the wave equation

$$\nabla^2\Psi - \frac{8\pi^2\mu V}{h^2}\Psi + \frac{4\pi\mu i}{h}\frac{\partial\Psi}{\partial t} = 0. \qquad (13\cdot1)$$

By hypothesis the corresponding probability of finding the particle in the element $dxdydz$ is $\Psi\Psi^*dxdydz$. Let us multiply this probability by the corresponding value of the coordinate x and sum over all elements of volume. Let the resulting integral be assumed to converge. It will clearly represent the statistical mean value of x for a large number of

[1] P. Ehrenfest, *Zeits. f. Physik* **45**, 455 (1927); A. E. Ruark, *J. Opt. Soc. Am.* **16**, 40 (1928), *Phys. Rev.* **31**, 533 (1928); A. Sommerfeld, *A.S.W.E.*, p. 290.

observations made on the various systems of an assemblage described by $\Psi(x, y, z, t)$. Calling this mean value $\bar{x}$, we have

$$\bar{x} = \int_\infty x\Psi\Psi^* dxdydz. \tag{13.2}$$

Similar expressions will hold for the other coordinates. In the limiting case of a very sharply defined wave packet, the expected departures of the individual measurements of x from the mean will be negligible so that $\bar{x}$ may be identified with the unique value assigned to x in the classical mechanics. Hence the equations of motion for the various mean values should become identical with the classical equations of motion in the case of such a sharply defined packet.

In particular, we should expect the classical or average momentum to be $\mu\dfrac{d\bar{x}}{dt}$. Computing the value of this quantity we obtain

$$\mu\frac{d\bar{x}}{dt} = \mu \int\int\int_\infty x\left(\Psi^*\frac{\partial\Psi}{\partial t} + \Psi\frac{\partial\Psi^*}{\partial t}\right)dxdydz. \tag{13.3}$$

The value of $\partial\Psi/\partial t$ is given by Eq. (13.1) and that of $\partial\Psi^*/\partial t$ by the conjugate equation

$$\nabla^2\Psi^* - \frac{8\pi^2\mu V}{h^2}\Psi^* - \frac{4\pi\mu i}{h}\frac{\partial\Psi^*}{\partial t} = 0. \tag{13.4}$$

Multiplying Eqs. (13.1) and (13.4) by $-\Psi^*$ and Ψ, respectively, and adding, we obtain

$$\Psi\nabla^2\Psi^* - \Psi^*\nabla^2\Psi - \frac{4\pi\mu i}{h}\left(\Psi\frac{\partial\Psi^*}{\partial t} + \Psi^*\frac{\partial\Psi}{\partial t}\right) = 0. \tag{13.5}$$

Hence Eq. (13.3) yields

$$\mu\frac{d\bar{x}}{dt} = \frac{h}{4\pi i}\int\int\int_\infty x \text{ div }(\Psi \text{ grad }\Psi^* - \Psi^* \text{ grad }\Psi)dxdydz$$

$$= \frac{h}{4\pi i}\int\int\int_\infty \text{ div }[x(\Psi \text{ grad }\Psi^* - \Psi^* \text{ grad }\Psi)]dxdydz$$

$$- \frac{h}{4\pi i}\int\int\int_\infty\left(\Psi\frac{\partial\Psi^*}{\partial x} - \Psi^*\frac{\partial\Psi}{\partial x}\right)dxdydz. \tag{13.6}$$

The first integral on the right may be converted into a surface integral with the aid of Gauss's theorem and this integral will vanish if Ψ approaches zero rapidly enough at infinity. We suppose this to be the case and thus obtain the equation

$$\mu\frac{d\bar{x}}{dt} = \frac{h}{4\pi i}\int\int\int_\infty\left(\Psi^*\frac{\partial\Psi}{\partial x} - \Psi\frac{\partial\Psi^*}{\partial x}\right)dxdydz. \tag{13.7}$$

The same assumption regarding the behavior of Ψ at infinity yields the relation

$$\int\int\int_\infty \left(\Psi^*\frac{\partial\Psi}{\partial x} + \Psi\frac{\partial\Psi^*}{\partial x}\right)dxdydz = \int\int\int_\infty \frac{\partial}{\partial x}(\Psi\Psi^*)dxdydz = 0.$$

Hence Eq. (13·7) reduces to the form

$$\mu\frac{d\bar{x}}{dt} = \frac{h}{2\pi i}\int\int\int_\infty \Psi^*\frac{\partial\Psi}{\partial x}dxdydz. \tag{13·8}$$

Differentiating again with respect to t and substituting the values of $\partial\Psi/\partial t$ and $\partial\Psi^*/\partial t$ from the wave equations (13·1) and (13·4), we obtain

$$\mu\frac{d^2\bar{x}}{dt^2} = \frac{h}{2\pi i}\int\int\int_\infty \left[\frac{\partial\Psi}{\partial x}\frac{\partial\Psi^*}{\partial t} + \Psi^*\frac{\partial}{\partial x}\left(\frac{\partial\Psi}{\partial t}\right)\right]dxdydz$$

$$= -\frac{h^2}{8\pi^2\mu}\int\int\int_\infty \left[\frac{\partial\Psi}{\partial x}\nabla^2\Psi^* - \Psi^*\nabla^2\left(\frac{\partial\Psi}{\partial x}\right) + \frac{8\pi^2\mu}{h^2}\Psi\Psi^*\frac{\partial V}{\partial x}\right]dxdydz.$$

Using Green's theorem and discarding the resulting surface integral as above, we obtain the relation

$$\mu\frac{d^2\bar{x}}{dt^2} = -\int\int\int_\infty \Psi\Psi^*\frac{\partial V}{\partial x}dxdydz \equiv -\overline{\frac{\partial V}{\partial x}}. \tag{13·9}$$

The right-hand member represents the mean value of the classical force in the x direction for the packet. As previously explained, in the limiting case of a sharply defined wave packet $\bar{x}(t)$ can be identified with the unique classical value of x at the time t and under the same circumstances we can identify $-\dfrac{\overline{\partial V}}{\partial x}$ with the unique classical value of the force component in the direction x at the time t. Thus the classical equations of motion are rigorously derived from the wave theory, subject only to possible restrictions regarding the way in which Ψ approaches zero at points far from the center of the packet. These restrictions are of little physical importance since it is certain that wave functions exist to which Eq. (13·9) applies, and which therefore are capable of representing the actual behavior of a large-scale particle. They are covered by the more general restrictions imposed on physically admissible wave functions in Sec. 32b.

14. THE STATISTICAL INTERPRETATION OF THE WAVE THEORY OF MATTER

14a. Review of Assumptions.—The mathematical developments of Secs. 9, 10, 11, and 12 will now be used to give a more precise form to the

statistical interpretation of our theory. To this end we may reformulate the physical assumptions made, either tacitly or in so many words, up to this point.

a. Matter, like radiation, combines the properties of the waves and corpuscles of classical physics. We may regard matter as corpuscular in character provided that we replace the classical laws of motion by statistical laws to be formulated in terms of a suitable "wave function" having properties similar in many respects to those of the wave fields of classical optics.

b. Whereas in the classical mechanics the "state" of a system is defined by the values of the coordinates and momenta of all its constituent particles, we here assume that the most exact description possible for our knowledge of the state of a system is contained in the specification of a corresponding function Ψ of the coordinates and the time which determines the relative probability of different positions or configurations for each value of the parameter t. In the following discussion we shall see that the Ψ function correlated with a system is not completely independent of the observer. Hence we shall say that all systems so prepared that they can be correlated with the same Ψ function are in a common *subjective state*. The *objective* or *true* state of an individual system is operationally undefined.

In case the probability of very large values of the coordinates is sufficiently small, the function Ψ is quadratically integrable and may be normalized in accordance with the condition

$$\int_\infty |\Psi|^2 dx_1 dy_1 \cdots dz_n = 1.$$

In this case $|\Psi|^2 dx_1 \cdots dz_n$ is to be interpreted as the probability of the group of mutually adjacent configurations associated with the volume element $dx_1 dy_1 \cdots dz_n$ in configuration space.[1]

c. The behavior of large-scale bodies as described by the Newtonian mechanics is to be correlated with the motion of appropriate wave packets or localized Ψ functions which may be thought of as compounded from elementary monochromatic and approximately plane[2] waves having a small continuous range of frequencies and wave normals.

d. The energy of a mechanical system and the frequency of the associated wave system are related according to the optical rule of Einstein,

$$E = h\nu. \tag{2·1}$$

---

[1] In view of later developments (*cf.* Secs. 19, and 36), it is best to give the phrase "probability of a group of configurations" operational meaning by identifying it with the probability that an accurate configuration measurement will yield a result which belongs to the group in question.

[2] *I.e.*, plane over the effective volume of the packet.

e. The vector local momentum of a single particle and the vector local wave number $\vec{\sigma}$ are related, in the limiting case where the classical mechanics is valid, according to the de Broglie rule,

$$\vec{p} = \hbar \vec{\sigma}. \tag{4.4}$$

This relation can be extended to the case of a system of n particles by a suitable definition of $\vec{p}$ and $\vec{\sigma}$ as vectors in the $3n$-dimensional configuration space.

f. The wave functions describing the behavior of an isolated system of particles moving under the influence of conservative forces are to be solutions of the Schrödinger wave equation (7.3), p. 22.

14b. Necessity of Introducing Assemblages.—It is necessary at this point to consider somewhat more narrowly the full import of these hypotheses. Let us begin with postulates **a** and **b**. The statistical laws and configuration probabilities referred to in these postulates presuppose the possibility of a large number of experiments in which the configuration and other properties of the system are observed. To say that the probability of a configuration range A is α means nothing, unless it is possible to make a large number of experimental determinations of the configuration. If these experiments are possible, the statement means, roughly speaking, that when their number is large the fraction of the whole number of cases in which the configuration lies in the range A is α.[1] A single simple observation can never test a law of probability and hence such laws can have meaning only in the description of the results of repeated experiments. Furthermore, if the probability of a configuration is a function of time, as we have supposed, the experiments or observations made to test the probability at any one time cannot be made on an individual mechanical system but require the use of an assemblage of similar independent systems which can be observed at corresponding times. In fact, even if the probability is independent of time, the assemblage is necessary owing to the fact that an observation involves an interaction between the system observed and the observing mechanism which inevitably modifies the future behavior of the system being studied. The configuration probabilities with which we normally deal are those of isolated systems or systems moving under the influence of known forces not including those used to make the observations. Hence the first observation we make upon a system initially in some

---

[1] The concept of probability has no precise operational meaning and is at the moment the subject of considerable controversy. Hence, from the operational point of view on which this volume is based, it would be better to avoid the term altogether. Such a procedure would involve some circumlocution, however, and it seems best for our present purpose to treat probability as an unanalyzable concept approximately defined in operational terms by the above statement.

definite subjective state will ordinarily throw it out of that state. Thus we are driven to the use of Gibbsian assemblages for the verification of the predictions of our wave functions.

When such an assemblage has been prepared, we can say that its state is definite and objective. Its reaction to future experiments will depend only on the nature of the experiments and not on the observers. If its past history has been such as to yield a maximum of information about the future, that history determines a wave function which predicts the statistical results of any future experiments on the assemblage and may be considered to define its state. In principle this wave function is experimentally determinable to a physically meaningless constant phase factor (*cf.* Secs. 15*f* and 42*b*). By mixing two or more such assemblages with different wave functions we get a more general type of assemblage called a mixture. We shall have more to say about mixtures in Secs. 41*b* and 53*d*. In the meantime we postulate that the most general possible assemblage of independent identical systems is a mixture whose state is defined when the wave functions and relative populations of its components are given.

When a physical system is a member of a Gibbsian assemblage previously prepared and having a definite wave function, it is natural to say that it is in the state defined by that wave function. More generally we can identify the state of a physical system with that of a Gibbsian assemblage of identical systems so prepared that the past histories (up to the epoch $t = 0$) of all its members are the same in all details that can affect future behavior as that of the original system. Such a definition of the state of an individual system is the best we can make but, as previously suggested, it contains a subjective element. Knowledge of the history of a system prior to a chosen epoch is necessarily subjective and can be different for different observers of that system. If two such observers A and B with different historical data try to repeat the preparation of the system in order to form a Gibbsian assemblage of systems in the same state, they will perform different operations and generate different types of assemblage. If each is a good physicist, each may deduce from his data a definite wave function describing the state of the system from his point of view and correctly predicting the behavior of the assemblage of similar systems which he has prepared. Now in exceptional cases where a wave function states that the probability of a certain experimental result is zero, the prediction can be disproved by a single experiment. In general, however, an experiment performed on a single system can neither prove nor disprove any wave function which we may attempt to correlate with it. Hence there is no general objective way in which a third person can prove that one of the two observers A and B is right and the other wrong. Thus it seems best to use the term "subjective state" in correlating a wave function with a specific system.

Whether we interpret the results derivable from the mathematical machinery of quantum mechanics in terms of previously prepared Gibbsian assemblages, or in terms of potential assemblages to be prepared according to our record of the past history of an individual system, we must interpret them in terms of some kind of assemblage. We are thus led to conceive of quantum mechanics as primarily a variety of statistical mechanics similar to the classical statistical mechanics of Gibbs.[1]

In the optical case it is proper to assume that the statistical behavior of the concrete assemblage of photons actually passing through a given piece of apparatus at a given time is the same as that of an ideal Gibbsian assemblage of photons correlated with wave systems of similar structure, but each moving in its own independent apparatus. This assumption is justified by the experimental fact that the behavior of the elements of the concrete assemblage is independent of the intensity of the light. In other words, the different photons of a beam of radiation have apparently complete mutual independence. Consequently simultaneous observations on the positions of a large number of photons made by allowing a beam of light to pass through dust-laden air, or to impinge on a photographic plate, may be regarded as the equivalent of a series of independent observations on the members of a corresponding Gibbsian assemblage of individual photons. Although closely packed electrons, protons, etc., are not independent, experience shows that in the case of low-density beams of such particles the interactions may be negligible, so that we can often replace independent observations on a Gibbsian assemblage with a single wholesale statistical observation of a concrete assemblage of approximately independent systems. Ideally, however, the predictions of quantum mechanics should be tested by a series of observations on a suitably prepared assemblage of completely independent systems each in its own separate box or laboratory.

---

[1] *Cf.* J. C. SLATER, *J. Franklin Inst.* **207**, 449 (1929).

In calling attention to the parallelism between quantum mechanics and classical statistical mechanics we must emphasize that the quantum-mechanical study of the behavior of assemblages whose statistical behavior can be described by a single wave function is *not* the quantum-mechanical way of treating the assemblages of the older form of statistical mechanics. In the latter we have infinitely many sharply defined classical states, whereas in the former we have a single quantum-mechanical state. The quantum-mechanical generalization of classical statistical mechanics, usually called *quantum statistical mechanics* (*cf.* Sec. 53), has to do with assemblages of systems including many quantum-mechanical states and thus involving many independent wave functions. We follow the notation of von Neumann (*M.G.Q.*, p. 158; see list of abbreviations, p. xviii) in calling such assemblages "mixtures" and will return to a brief discussion of their properties in Sec. 41b. In this chapter the discussion refers throughout to the simpler type of assemblage, sometimes called a "pure case," all members of which belong to a common subjective state with a common single wave function.

14c. Multiplicity of Energy and Momentum Values for a Definite State.—Having settled this point, we turn our attention to postulate **d** which may conceivably be interpreted in two ways. Since the frequency ν does not have a unique value for any but the most special wave functions, we must either interpret Eq. (2·1) as indicating an indeterminacy in the energy of a particle associated with a wave packet, or else we must regard the value of ν which appears in this equation as a mean value for the wave function as a whole. We choose the former alternative as the more natural and logical. It is more natural since we are endeavoring to construct a mechanical theory which parallels the dualistic theory of radiation, and since it is well known that the photons associated with a given electromagnetic wave system may have a variety of energies. It is more logical, for it can be proved a consequence of postulate **b** combined with a reasonable experimental definition of kinetic energy.

To verify this last statement let us consider the case of a wave function composed of the sum of two disturbances Ψ_1 and Ψ_2, each of which forms a typical wave packet with a fairly well defined frequency and direction of motion.[1] In the optical case such a disturbance can be obtained experimentally by allowing a beam of plane-parallel monochromatic radiation to fall upon an aperture covered by a shutter. Opening and closing the shutter momentarily would form a primary packet which could be split into two parts with different wave lengths and directions of motion by suitable partial reflection at the surface of a moving half-silvered mirror. If a disturbance of this kind is associated with a single particle, whether electron or photon, hypothesis **b** requires that the particle have a certain probability of moving with packet 1 and a complementary probability of moving with packet 2. If an experimental observation showed the particle to be in packet 1, we should have to assign to it one energy and momentum, while if it were found in packet 2, we should have to assign to it a different energy and momentum. Hence, in this case the correlation of a single energy and single momentum with the wave function leads to an absurd result; and, since no sharp dividing line can be drawn between this case and that of a typical single wave packet, we conclude that unique energy and momentum values should not be assigned to any wave packet.

Indeed, we know by direct experiment in the optical case that such a single packet always contains a range of possible energies and momenta. If the packet be formed with the aid of an aperture and shutter as suggested above, diffraction will cause the radiation to diverge by an amount varying inversely with the dimensions of the aperture. If the intensity is large and the aperture is small compared with the wave length, photons will proceed in all directions from the slit, carrying

[1] As the Schrödinger equation (7·3) is linear and homogeneous, the sum of any two solutions is a solution.

with them momenta directed along the radius vector from the slit. At the same time the interruption of the primary beam by the shutter will destroy its monochromatic character and scatter the photons over a narrow continuous spectrum. Diffraction experiments involving prolonged photographic exposures and very low intensities show that the distribution of energy over the pattern is independent of the intensity. We infer that, if the intensity of the radiation in the packet is so low that only one or two photons pass through the aperture during its formation, the relative probability of each energy and direction of motion is the same as if the intensity were very large.

The above discussion leads naturally to a reinterpretation or extension of postulate **e** which relates the classical local momentum

$$\sqrt{2\mu[E - \overline{V(x,y,z)}]}$$

with the local wave number $\sigma = 1/\lambda = \nu/w(x,y,z,\nu)$. These concepts apply, strictly speaking, to classical particles and monochromatic waves only. In the case of an assemblage of particles having a variety of energies, the concept of local momentum becomes indeterminate, while the local wave number becomes equally indefinite when applied to the wave packet which represents such an assemblage.

To sharpen our ideas we shall substitute the idea of a *measured momentum* for that of classical local momentum and the concept of *sinusoidal wave number and wave length* for local wave number and wave length. Thus we shall suppose that a wave system in one dimension has a definite wave length if, and only if, it is strictly sinusoidal. Similarly a wave system in three dimensions will be said to have a definite vector wave number $\overrightarrow{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, only if it has the special sinusoidal plane wave form

$$\Psi = F(t)e^{2\pi i(x\sigma_x + y\sigma_y + z\sigma_z)}.$$

As wave packets are in general representable by Fourier integrals, we must regard them as compounds of waves having infinitely many wave lengths and wave numbers just as they have infinitely many frequencies.

Furthermore we shall suppose that an assemblage of similarly prepared particles has a single definite momentum, only if a measurement of the momentum of each particle would give a single unique result in all cases. As indicated above, such a measurement would lead in general to a wide variety of results and so we must consider that the assemblage consists of particles having a great many different momenta.

Now in the limiting case of an assemblage of heavy particles prepared in such a manner as to minimize the uncertainties in position, energy, and momentum, the values of momentum obtained by measurement would converge upon the value of the classical local momentum cor-

responding to the mean position and energy of the particles in the assemblage. Similarly the wave lengths and wave numbers obtained by Fourier analysis of the wave packet which represents such an assemblage will converge on the local wave length for the average frequency of the packet at the center of the packet. Hence we infer that in this limiting case we can use the equation

$$\vec{p} = h\vec{\sigma}$$

to correlate values of the experimental vector momentum and the vector wave number obtained by Fourier analysis. In the next section it will be proved that if we make a suitable experimental definition of measured momentum, we can use the equation in the same way quite generally.

15. THE WAVE FUNCTION AND MEASUREMENTS OF LINEAR MOMENTUM

15a. Operational Definition of Momentum for Free Particles.—In the last section we introduced the postulate that our knowledge of the state of a system cannot be more exactly specified than by a function $\Psi(x,t)$ which determines the probability of various configurations as a function of time and which describes the statistical behavior of an assemblage of identical independent systems all of which are in the same subjective state. We saw further that, just as the systems of such an assemblage have a multiplicity of possible positions, so we must think of them as having a similar multiplicity of possible momenta.

The question now arises: How can we determine from the wave function of an assemblage of atoms or particles what the probabilities of various possible momenta are? The answer to this question must, of course, agree with experiment and must consequently have its basis in an experimental definition of momentum. The first step toward the answer to the above question is therefore to agree on an experimental procedure for the measurement of linear momentum. Whatever procedure we accept becomes the "operational"[1] definition of momentum for quantum mechanics. It must harmonize with the classical definition in the limiting case of a sharply defined wave packet and must be so formulated that the measurement is feasible even when the classical mechanics is inapplicable.

A considerable variety of classical methods is available for the measurement of momentum. Of these perhaps the most fundamental are the magnetic deflection method involving a separate measurement of the charge on the particle, and the elementary method of measuring the velocity directly and multiplying it by the separately measured mass. The fact that these measurements are not complete in one operation is an objectionable feature apparently shared in one way or another by

---

[1] *Cf.* P. W. Bridgman, *The Logic of Modern Physics*, New York, 1927.

every other method. As the mass and charge are parameters which go into the construction of the Schrödinger equation, the necessity of measuring them separately is perhaps less disturbing than otherwise. We here adopt the direct velocity type measurement as the most fundamental procedure and the one best adapted to our purpose. Thus our starting point is the formula

$$\vec{p} = \mu\vec{v}. \tag{15·1}$$

In order to measure the velocity $\vec{v}$ directly it is necessary to make two observations of the positional vector $\vec{r}$ and apply the classical formula for the average velocity, *viz.*,

$$\vec{\bar{v}} = \frac{\vec{r_2} - \vec{r_1}}{t_2 - t_1}. \tag{15·2}$$

If the particle under consideration is moving in a field-free space, we can identify the average velocity with the instantaneous velocity and use (15·1) and (15·2) to reduce the problem of measuring momentum to that of making two successive measurements of position. We encounter an immediate difficulty, however, in the fact that every measurement of position involves an unpredictable change in momentum. A determination of position is usually made by means of radiation with the aid of a microscope, telescope, or slit system. It depends upon a collision phenomenon, *viz.*, the scattering or reflection of photons by the particle or system under observation. The study of the recoil electrons produced in the scattering of X-rays by light atoms (Compton scattering) shows that in such a collision the scattering system undergoes a change of momentum of the order of h/λ for each photon scattered. This change in the momentum of the system, whose position is being observed, can be made very small by using low-intensity, long wave length radiation, but unfortunately the diffraction of the radiation causes an uncertainty in the positional measurement which cannot be reduced below a value of the order of magnitude of the wave length λ. Hence the momentum perturbation is directly proportional to the precision of the observation of position and cannot be neglected in the exact location of an atomic system. Other methods of measuring position can be shown to involve the same difficulty (*cf.* Sec. 16).

It will be observed that the change in momentum involved in the first measurement of position is the one that makes the difficulty. We take the object of the measurement to be the determination of the momentum immediately prior to the time t_1 when the measurement is begun. As Eq. (15·2) gives the average velocity in the time interval

$t_1 < t < t_2$, a momentum perturbation at the time t_2 can do no harm. On the other hand, the result of the initial perturbation is that the momentum we actually measure differs from that which we wish to measure by an amount Δp whose absolute value is of the order of $h/\Delta r$, where Δr is the absolute value of the uncertainty in $\vec{r}_1$. Fortunately, if the particle is in an unlimited field-free space, as we shall assume for the present, we can make the initial observation of position a very rough one and still get accurate values of $\vec{v}$ by allowing the time interval $t_2 - t_1$ to be very large. *In fact the accuracy of the momentum observations can be pushed beyond any assignable limit if we start with an assemblage of particles whose initial position has already been roughly measured and make $t_2 - t_1$ large enough.* Actually the assumption that the position of the particles of our assemblage has been roughly measured prior to the measurement of momentum is almost no restriction at all, for our whole discussion has to do with observations on an assemblage of particles whose state is described by a Ψ function which can be normalized so that we can interpret $|\Psi|^2 d\tau$ as the probability of the positions in the volume element $d\tau$. It is but a slight step farther to require that our wave function shall approach zero rapidly enough at infinity so that the integrals which give the average value of each of the coordinates x, y, z and the probable errors $[\overline{(x - \bar{x})^2}]^{1/2}$, $[\overline{(y - \bar{y})^2}]^{1/2}$, $[\overline{(z - \bar{z})^2}]^{1/2}$ shall exist. If the wave function is of this type, it determines the initial position of the particle to a certain degree of approximation and we can infer that the initial position is determined to the same degree of approximation by the experimental procedure for preparing the particles which make up the assemblage. Obviously an assemblage of identical particles in a common subjective state having the wave function Ψ can be prepared only if there is an experimental procedure for assuring that no particle can get into the assemblage which does not have the right wave function. This procedure would accordingly constitute a rough measurement of position made prior to the measurement of momentum.

It will now be evident that in the case of an assemblage of identical particles prepared so as to have a roughly known initial position and moving in an unlimited field-free space (the apparatus used for their preparation having been withdrawn) Eqs. (15·1) and (15·2) afford the basis for momentum measurements which can be carried through, in principle, to any desired degree of exactitude. We now address ourselves to the task of computing from the initial wave function the probabilities of different possible linear momenta for an assemblage of this type, reserving for later discussion the case in which the particles are moving in a force field.

15b. Computation of Momentum Probabilities from a Wave Function. Taking the origin at or near the most probable initial position, we can

replace Eqs. (15·1) and (15·2) by the relation

$$\vec{p} = \frac{\mu}{t}\vec{r},\qquad(15\cdot3)$$

in which r is the measured position at some time t, great enough so that r is large compared with the uncertainty in its measurement. Then to a certain degree of approximation we may identify the probability of a momentum vector $\vec{p}$ whose terminal point lies within the element of volume $dp_x dp_y dp_z$ of momentum space[1] with the probability of locating the particle itself in the corresponding element of ordinary space, *viz.*, the element $dxdydz = (t/\mu)^3 dp_x dp_y dp_z$, for which $\vec{r} = t\vec{p}/\mu$. As t is made larger and larger, the precision of the measurement improves, so that we can rigorously identify the limit of this probability with the probability of the corresponding measured momentum.

As a basis for the computation of the probability we use the familiar principle of physical optics, that a continuous distribution of elementary wavelets (Huygens' wavelets, or plane waves) destroys itself by mutual interference except at points where there is complete or partial phase agreement among its elements. This principle of phase agreement was used in Sec. 12 to demonstrate the laws of motion of wave packets.

In order to apply the principle to the problem in hand, we may suppose the wave packet analyzed into a continuous plane-wave spectrum such as that given by Eq. (10·3). Then, at any space-time point x,y,z,t, we shall have phase agreement for the elementary range of wave numbers $d\sigma_x d\sigma_y d\sigma_z$, provided that the differential of the phase function φ [*cf.* Eq. (10·5)] vanishes for the range in question. In other words, we must have

$$d\varphi = 2\pi\left[\left(x - t\frac{\partial\nu}{\partial\sigma_x}\right)d\sigma_x + \left(y - t\frac{\partial\nu}{\partial\sigma_y}\right)d\sigma_y + \left(z - t\frac{\partial\nu}{\partial\sigma_z}\right)d\sigma_z\right] = 0.$$

Thus it follows from the principle of phase agreement that the value of the wave function at the space-time point x,y,z,t depends almost exclusively on the amplitudes of the plane-wave components in the neighborhood of the one for which

$$x = t\frac{\partial\nu}{\partial\sigma_x},\qquad y = t\frac{\partial\nu}{\partial\sigma_y},\qquad z = t\frac{\partial\nu}{\partial\sigma_z}.\qquad(15\cdot4)$$

In the case of free particles, ν is given by Eq. (10·2). Introducing this value into Eqs. (15·4) and solving for $h\sigma_x$, $h\sigma_y$, and $h\sigma_z$, we obtain

$$h\sigma_x = \frac{\mu x}{t},\qquad h\sigma_y = \frac{\mu y}{t},\qquad h\sigma_z = \frac{\mu z}{t}.\qquad(15\cdot5)$$

---

[1] A space in which the momentum vectors are laid off as displacement vectors from the origin of a set of Cartesian coordinate axes.

It will be proved that in the limit of very large values of t the wave function depends solely on the amplitude of the plane-wave component for which the above equations are satisfied. The right-hand member of each of these equations denotes one of the momentum components to be assigned to the point x,y,z,t according to the definition of Eq. (15·3), while the left-hand member is the value of the same momentum component given by the basic postulate of Eq. (4·4) (p. 14). We conclude that the operational definition of momentum is in accord with Eq. (4·4) and permits us to interpret this equation as a relation between measured momenta and the wave numbers of Fourier analysis.

Let us now proceed to the rigorous justification of the statements in the preceding paragraph and to the exact evaluation of the probability of the range of momentum values $dp_x dp_y dp_z$. The first step is to rewrite Eq. (10·3) in the form

$$\Psi(x,y,z,t) = \int\int\int_\infty G(\sigma_x,\sigma_y,\sigma_z)e^{2\pi i\left[x\sigma_x + y\sigma_y + z\sigma_z - \frac{ht}{2\mu}(\sigma_x{}^2 + \sigma_y{}^2 + \sigma_z{}^2)\right]}d\sigma_x d\sigma_y d\sigma_z. \quad (15\cdot6)$$

By an appropriate change of variables and integration by parts this reduces to (cf. Appendix B)

$$\Psi(x,y,z,t) = -\left(\frac{\mu}{2ht}\right)^{3/2} 2(i + 1)e^{\frac{\pi i\mu}{ht}(x^2+y^2+z^2)}G\left(\frac{\mu x}{ht}, \frac{\mu y}{ht}, \frac{\mu z}{ht}\right) + \text{negligible}$$

$$\text{terms in } \left(\frac{\mu}{2ht}\right)^2. \quad (15\cdot7)$$

This proves that for very large values of t the wave function depends solely on the value of the amplitude function G for that point $\sigma_x,\sigma_y,\sigma_z$ which satisfies Eqs. (15·5). To get the probability dW of a momentum in the element $dp_x dp_y dp_z$ of momentum space we have to evaluate $\Psi\Psi^*$ for the point

$$x = \frac{tp_x}{\mu}, \qquad y = \frac{tp_y}{\mu}, \qquad z = \frac{tp_z}{\mu},$$

[cf. Eq. (15·3)], multiply it by the magnitude of the corresponding volume element $(t/\mu)^3 dp_x dp_y dp_z$, and then pass to the limit of infinite values of t.[1] We obtain

$$dW = \frac{1}{h^3}G\left(\frac{p_x}{h}, \frac{p_y}{h}, \frac{p_z}{h}\right)G^*\left(\frac{p_x}{h}, \cdot \cdot \right)dp_x dp_y dp_z.$$

It is convenient at this point to change the variables of integration in Eq. (15·6) from σ_x, σ_y, σ_z to p_x, p_y, p_z in accordance with the universal relation $\vec{p} = h\vec{\sigma}$. The result is

$$\Psi(x,y,z,t) = \frac{1}{h^3}\int\int\int_\infty G\left(\frac{p_x}{h}, \frac{p_y}{h}, \frac{p_z}{h}\right)e^{\frac{2\pi i}{h}\left[(xp_x + yp_y + zp_z) - \frac{t}{2\mu}(p_x{}^2 + p_y{}^2 + p_z{}^2)\right]}dp_x dp_y dp_z.$$

$$(15\cdot8)$$

---

[1] The wave function Ψ is assumed to be normalized.

The introduction of the function

$$\Phi(p_x,p_y,p_z,t) \equiv h^{-3/2}G\left(\frac{p_x}{h}, \frac{p_y}{h}, \frac{p_z}{h}\right)e^{-\frac{\pi i t}{h\mu}(p_x{}^2+p_y{}^2+p_z{}^2)}$$

$$= h^{-3/2}G\left(\frac{p_x}{h}, \cdot \cdot \cdot\right)e^{-2\pi i\nu t}$$

throws Eq. (15·8) into the form

$$\Psi(x,y,z,t) = h^{-3/2}\iiint_\infty \Phi(p_x,p_y,p_z,t)e^{\frac{2\pi i}{h}(xp_x + yp_y + zp_z)}dp_xdp_ydp_z. \quad (15\cdot9)$$

The Fourier integral theorem then yields the symmetrical reciprocal formula

$$\Phi(p_x,p_y,p_z,t) = h^{-3/2}\iiint_\infty \Psi(x,y,z,t)e^{-\frac{2\pi i}{h}(xp_x + yp_y + zp_z)}dxdydz \quad (15\cdot10)$$

[*cf.* Eq. (9·5)], and the expression for the probability of the momentum range $dp_xdp_ydp_z$ reduces to

$$dW = \Phi\Phi^*dp_xdp_ydp_z \quad (15\cdot11)$$

In view of Plancherel's theorem we may deduce from Eqs. (15·9) and (15·10) the additional relation

$$\iiint_\infty \Phi\Phi^*dp_xdp_ydp_z = \iiint_\infty \Psi\Psi^*dxdydz = 1 \quad (15\cdot12)$$

which is to be interpreted as a statement that the sum of the probabilities of all possible momenta is unity, as it should be.

Since Φ plays the same role in determining momentum as Ψ plays in determining position, it is convenient to introduce Jordan's nomenclature and designate Φ and Ψ as the *probability amplitudes* for momentum and position, respectively. The appropriateness of this nomenclature is evident from Eqs. (8·1) and (15·11).

The above Eqs. (15·10), (15·11), and (15·12) complete the development of the mathematical method for evaluating the relative probabilities of different momenta in the case of a free particle.

*15c. Momentum of Center of Gravity.—The method is readily extended to the problem of the momentum associated with motion of the center of gravity of a system of n particles moving under the influence of a potential-energy function which depends only on the relative coordinates of the system. In this case we define the momentum as the product of the total mass M and the vector velocity of the center of gravity. We assume as before that the initial position is approximately determined by the method of starting the system off. The velocity is then determined by a single observation of position made a long while after the starting time. In the limit there is a one-to-one correspondence between the positional coordinates observed and the momentum values to be assigned, so that the observation of position and of momentum are equivalent.

In order to compute the expectation of any given range of momentum values it is necessary to make a change of independent variables so as to express the wave function in terms of the coordinates of the center of gravity and a suitable set of relative coordinates for the particles which make up the system.   Let the absolute coordinates of the n particles be $x_1,\ y_1,\ z_1,\ x_2,\ \cdot\ \cdot\ \cdot\ ,\ z_n$ and let $X,\ Y,\ Z$ be the coordinates of the center of gravity so that

$$MX = \sum_{k=1}^{n}\mu_k x_k, \qquad MY = \sum_{k=1}^{n}\mu_k y_k, \qquad MZ = \sum_{k=1}^{n}\mu_k z_k. \quad (15\cdot13)$$

Let the relative coordinates be

$$\left. \begin{aligned} \xi_k &= x_k - x_1, \\ \eta_k &= y_k - y_1, \\ \zeta_k &= z_k - z_1. \end{aligned} \right\} \qquad k = 2,\ 3,\ \cdot\ \cdot\ \cdot\ ,\ n. \qquad (15\cdot14)$$

Substitution of the new variables in the wave function $\Psi_0(x_1,\ \cdot\ \cdot\ \cdot\ ,\ z_n,t)$ transforms it into a function of the new variables which we shall call $\Psi(X,Y,Z,\xi_2,\eta_2,\cdot\ \cdot\ \cdot\ ,\zeta_n,t)$.   Thus

$$\Psi_0(x_1,\ \cdot\ \cdot\ \cdot\ ,z_n,t) \to \Psi(X,Y,Z,\xi_2,\ \cdot\ \cdot\ \cdot\ ,\zeta_n,t).$$

The Jacobian, or functional determinant, of the transformation is unity, so that the volume element $dx_1\ \cdot\ \cdot\ \cdot\ dz_n$ goes over into the volume element $dXdY\ \cdot\ \cdot\ \cdot\ d\zeta_n$.   Since $\Psi_0\Psi_0^* dx_1\ \cdot\ \cdot\ \cdot\ dz_n$ is the probability of a configuration in the element $dx_1\ \cdot\ \cdot\ \cdot\ dz_n$, we may identify

$$\Psi\Psi^* dX\ \cdot\ \cdot\ \cdot\ d\zeta_n$$

with the probability of a configuration in the element $dX\ \cdot\ \cdot\ \cdot\ d\zeta_n$.

The wave equation expressed in terms of the new variables is obtained by direct transformation and has the form[1]

$$\frac{1}{M}\left(\frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}\right)\Psi + \frac{1}{\mu_1}\sum_{j=2}^{k}\sum_{k=2}^{n}\left(\frac{\partial^2}{\partial\xi_j\partial\xi_k} + \frac{\partial^2}{\partial\eta_j\partial\eta_k} + \frac{\partial^2}{\partial\zeta_j\partial\zeta_k}\right)\Psi$$

$$+ \sum_{k=2}^{n}\frac{1}{\mu_k}\left(\frac{\partial^2}{\partial\xi_k^2} + \frac{\partial^2}{\partial\eta_k^2} + \frac{\partial^2}{\partial\zeta_k^2}\right)\Psi - \frac{8\pi^2 V(\xi_2,\ \cdot\ \cdot\ \cdot\ \zeta_n)}{h^2}\Psi$$

$$+ \frac{4\pi i}{h}\frac{\partial\Psi}{\partial t} = 0. \quad (15\cdot15)$$

---

[1] The cross-derivatives can be eliminated if we use as relative coordinates the quantities

$$\xi_k = x_k - \sum_{j=1}^{k-1}\mu_j x_j \Big/ \sum_{j=1}^{k-1}\mu_j, \qquad \eta_k = \cdot\ \cdot\ \cdot\ .$$

This elimination, however, somewhat complicates the expression for the potential energy.

Special solutions of this equation may be found by the device of separating the variables.   One assumes that Ψ has the form

$$\Psi = \Psi_g(X,Y,Z,t)\Psi_r(\xi_2, \cdots \zeta_n,t). \tag{15.16}$$

Insertion of this expression into Eq. (15.15) shows that it gives a solution provided that Ψ_g and Ψ_r satisfy

$$\frac{1}{M}\left(\frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}\right)\Psi_g + \frac{4\pi i}{h}\frac{\partial \Psi_g}{\partial t} = 0 \tag{15.17}$$

and

$$\sum_{k=2}^{n}\frac{1}{\mu_k}\left(\frac{\partial^2}{\partial \xi_k^2} + \frac{\partial^2}{\partial \eta_k^2} + \frac{\partial^2}{\partial \zeta_k^2}\right)\Psi_r + \frac{1}{\mu_1}\sum_{j=2}^{k}\sum_{k=2}^{n}\left(\frac{\partial^2}{\partial \xi_j \partial \xi_k} + \frac{\partial^2}{\partial \eta_j \partial \eta_k}\right.$$
$$\left. + \frac{\partial^2}{\partial \zeta_j \partial \zeta_k}\right)\Psi_r - \frac{8\pi^2 V}{h^2}\Psi_r + \frac{4\pi i}{h}\frac{\partial \Psi_r}{\partial t} = 0, \tag{15.18}$$

respectively.   If the wave function factors in this manner into the product of a function of the coordinates of the center of gravity and a function of the relative coordinates, the probability of any set values for X, Y, Z is independent of the values of the other variables.

More general solutions of Eq. (15.15) can be obtained by compounding different special solutions of Eqs. (15.17) and (15.18).   This fact suggests the possibility of expressing any quadratically integrable solution of Eq. (15.15) in the form

$$\Psi = \frac{1}{h^{3/2}}\int\int\int_{\infty} Q(P_x,P_y,P_z,\xi_2, \cdots \zeta_n,t)e^{\frac{2\pi i}{h}[(XP_x + YP_y + ZP_z) - E_g t]}dP_x dP_y dP_z, \tag{15.19}$$

in which Q is a solution of Eq. (15.18) involving P_x, P_y, P_z as parameters, while E_g is defined by

$$E_g = \frac{P_x^2 + P_y^2 + P_z^2}{2M} \tag{15.20}$$

so as to make the exponential factor in the integrand of Eq. (15.19) a plane-wave solution of Eq. (15.17).   The conjecture is readily verified,[1]

---

[1] Given an arbitrary quadratically integrable solution of Eq. (15.15), we may always apply the Fourier integral theorem to the variables X, Y, Z and so express the function in the form

$$\Psi = h^{-3/2}\int\int\int_{\infty} G e^{\frac{2\pi i}{h}(XP_x + YP_y + ZP_z)}dP_x dP_y dP_z.$$

If we introduce the function

$$Q = G e^{\frac{2\pi i}{h}E_g t},$$

Ψ takes the form of Eq. (15.19).   Then substitution in the wave equation (15.15) shows that Q must be a solution of Eq. (15.18).

and hence the expression (15·19) can be used for the evaluation of the probabilities of different possible momentum ranges.

Now let us suppose that we observe the configuration of the system for some very large value of the time t. As in the case of a single free particle, the vector momentum $\vec{P}$ is to be determined from the positional coordinates X, Y, Z in accordance with the relation

$$\vec{P} = \frac{M}{t}\vec{R},$$

in which $\vec{R}$ is the position vector for the center of gravity (X,Y,Z). Applying the theorem of Eq. (15·7) we find that, in the limit of large values of t, $\Psi(X,Y,Z,\xi_2, \cdots \zeta_n,t)$ is determined by $Q(P_x,P_y,P_z,\xi_2, \cdots \zeta_n,t)$ with

$$P_x = \frac{MX}{t}, \qquad P_y = \frac{MY}{t}, \qquad P_z = \frac{MZ}{t}.$$

Hence the parameters P_x, P_y, P_z of Eq. (15·19) are to be identified with the components of $\vec{P}$. The probability that the variables P_x, P_y, P_z, ξ_2, $\cdots \zeta_n$ lie in the elementary range $dP_x dP_y dP_z d\xi_2 \cdots d\zeta_n$ is then

$$dW = \Psi\Psi^* \frac{t^3}{M^3} dP_x dP_y dP_z d\xi_2 \cdots d\zeta_n$$

$$= QQ^* dP_x dP_y dP_z d\xi_2 \cdots d\zeta_n. \tag{15·21}$$

Let dW_t denote the total probability of the momentum range

$$dP_x dP_y dP_z$$

for all values of the relative coordinates. This quantity is obtained by integrating the expression for dW as follows:

$$dW_t = dP_x dP_y dP_z \int_\infty QQ^* d\xi_2 \cdots d\zeta_n. \tag{15·22}$$

Similarly the probability that the individual momentum component P_x will lie in the range $P_x' < P_x < P_x' + dP_x$ is

$$dP_x \int QQ^* dP_y dP_z d\xi_2 \cdots d\zeta_n \tag{15·23}$$

*15d. The Measurement of the Momenta of Particles Moving in a Force Field.—In the case of particles moving in a force field the above scheme of measurement breaks down. The velocity cannot be treated as constant, and its instantaneous value can be closely approximated by means of Eq. (15·2) only if (a) the time interval $t_2 - t_1$ is made short, and (b) the two positional measurements are made with great accuracy. We cannot in general employ the initial positional measurement involved in the preparation of the particles, for that will not always be accurate enough. Hence we again meet the difficulty due to the perturbation of the momentum by the initial positional measurement.

One conceivable way of avoiding this difficulty would be to measure the velocity of the particle by means of the Doppler effect which it imparts to long wave length reflected light. This procedure fails to work out, however, because the frequency of light waves (energy of photons) can be measured accurately only if we have to do with a long train of equally spaced waves—a condition not fulfilled by the reflected light in this case.

It therefore seems necessary to base our definition of momentum on an artificial and apparently impossible experiment in which the observer first *abolishes the force field and then measures the momentum* at his leisure by the method previously worked out for free particles! Although we cannot actually remove a force field instantaneously, we can in many cases remove it so rapidly that the hypothetical experiment described may be regarded as a limiting case of a procedure which can be realized in practice. Hence this method of observation is in the same class with all the idealized experiments used to give operational[1] meaning to definitions in physics. In particular, we may note that the exact measurement of the position of a particle is also impossible in practice, and that we can give position an exact operational meaning only by postulating a limiting experiment with radiation of infinitely short wave length.

Now it follows from the wave equation that $\partial\Psi/\partial t$ remains finite so long as the potential energy is finite. Consequently the change in Ψ during the period of removal of the force field approaches zero as the length of that period approaches zero. Hence, if we define the instantaneous momentum as the momentum obtained by suddenly removing the force field and carrying out a measurement by the method used for free particles, the probability of any particular value will be related to the instantaneous form of the wave function exactly as if the particle were free. In other words, the probability of the momentum range $dp_x dp_y dp_z$ is given as before by Eq. (15·11) with Φ, the complex Fourier amplitude for the momentum $\vec{p}$, given by Eqs. (15·9) and (15·10). The only difference between the two cases is that, if the particle is not free, Φ is not a harmonic function of t and hence $\Phi\Phi^*$ is not constant in time.[2]

---

[1] See footnote 1, p. 58.

[2] In Sec. 12 we showed that the vector group velocity $\vec{v}_{g}$ of a symmetrical wave packet taken as a whole is equal to $\hbar\vec{\sigma}_0/\mu$, where σ_0 is the wave number with maximum amplitude obtained by an instantaneous Fourier analysis of the packet into plane-wave components. $\hbar\vec{\sigma}_0$ is then identical with the most probable momentum given by the method of observation described above. It is very satisfactory to find that the most probable momentum is equal to $\mu\vec{v}_g$.

If the force field has a vector potential $\vec{a}$(*cf.* Sec. 7), the momentum is defined not by Eq. (15·1), but by the relation

$$\vec{p} = \mu\vec{v} + \frac{e}{c}\vec{a}. \tag{15·24}$$

If we suddenly reduce $\vec{a}$ to zero, the induced electric force given by

$$\vec{\mathcal{E}} = -\frac{1}{c}\frac{\partial\vec{a}}{\partial t}$$

would in the classical theory give an impulse to the particle just equal and opposite to the change in $\frac{e}{c}\vec{a}$. Thus $\vec{p}$ is really unaffected by suddenly reducing $\vec{a}$ to zero. The wave function is also unchanged and we conclude that the probability of different momentum values is given by Eqs. (15·10) and (15·11) as before.

*15e. Individual Momenta of Particles in a System.—Finally, let us consider the determination of the individual momenta of the particles in a system moving under the influence of a potential energy V. For simplicity we shall suppose that there are only two particles in the system, the extension of the argument to more general cases being relatively simple. In order to permit a measurement we shall suppose that, at an arbitrary instant t_0, the potential function and the mutual forces between the particles are abolished. We then observe the positions of the two particles after a long period of time and compute the momenta as before. For values of t greater than t_0 the wave function $\Psi(x_1,y_1,z_1,x_2,y_2,z_2,t)$ will then obey the equation

$$\frac{1}{\mu_1}\left(\frac{\partial^2}{\partial x_1{}^2} + \frac{\partial^2}{\partial y_1{}^2} + \frac{\partial^2}{\partial z_1{}^2}\right)\Psi + \frac{1}{\mu_2}\left(\frac{\partial^2}{\partial x_2{}^2} + \frac{\partial^2}{\partial y_2{}^2} + \frac{\partial^2}{\partial z_2{}^2}\right)\Psi + \frac{4\pi i}{h}\frac{\partial\Psi}{\partial t} = 0. \tag{15·25}$$

Any quadratically integrable solution of this equation can be expressed in the form

$$\Psi = \frac{1}{h^3}\int \cdots \int_{\infty} G(\alpha_1 \cdots \gamma_2)e^{-\frac{2\pi i E(t-t_0)}{h}}e^{\frac{2\pi i}{h}(x_1\alpha_1 + \cdots + z_2\gamma_2)}d\alpha_1 \cdots d\gamma_2,$$

where α_1, β_1, γ_1 and α_2, β_2, γ_2 are to be identified with the momentum components of particles 1 and 2 respectively, while

$$E = \frac{\alpha_1{}^2 + \beta_1{}^2 + \gamma_1{}^2}{2\mu_1} + \frac{\alpha_2{}^2 + \beta_2{}^2 + \gamma_2{}^2}{2\mu_2}.$$

For very large values of $t - t_0$ we can prove, in analogy with Eq. (15·7), that

$$\Psi \cong \left[\frac{\mu_1}{2h(t - t_0)}\right]^{3_2}\left[\frac{\mu_2}{2h(t - t_0)}\right]^{3_2} 8iG\left(\frac{\mu_1 x_1}{h(t - t_0)}, \frac{\mu_1 y_1}{h(t - t_0)}, \cdots\right) \times$$
$$e^{\frac{\pi i}{h(t - t_0)}[\mu_1(x_1^2 + y_1^2 + z_1^2) + \mu_2(x_2^2 + y_2^2 + z_2^2)]}$$

It follows that the probability of the elementary momentum range $d\alpha_1 d\beta_1 d\gamma_1 d\alpha_2 d\beta_2 d\gamma_2$ is

$$dW = GG^* d\alpha_1 \cdots d\gamma_2 = \Phi\Phi^* d\alpha_1 \cdots d\gamma_2$$

with

$$\Phi = Ge^{-\frac{2\pi i E(t - t_0)}{h}}$$

To get the probability of any individual momentum range $d\alpha_1$, we have only to integrate over all values of the other variables.

15f. Summary: The Determination of the Wave Function of an Assemblage.— In concluding this discussion of measured linear momentum the following points are to be emphasized:

a. Linear momentum has been defined by means of an idealized experiment which is in agreement with the classical conception. According to this definition the momentum is determined by a measurement of mass and by two positional observations in a field-free space separated by a large time interval. The first of these two observations may be identified with the process of preparing the assemblage of systems under consideration which will normally locate the members more or less exactly in space. The large time interval is introduced because the positional observations are inexact.[1]

b. By hypothesis the behavior in time of an assemblage of systems with a common initial subjective state is determined by a wave function Ψ which obeys the Schrödinger equation (7·3) and whose initial form is fixed by experimental conditions involved in the initial preparation of the assemblage. We have computed the probability of the various momentum values in terms of the instantaneous form of this function. In the case of a free particle, the momentum probability is independent of time.

c. As regards relations between the measured momentum and the classical local momentum, we may observe that in the limiting case of a sharply defined wave packet to which the classical mechanics is appli-

---

[1] Strictly speaking, we should note that a positional observation is a correlation of a pair of values of positional coordinate and time in which there may be as much uncertainty about the time as about the positional coordinate. Uncertainties in time as well as position are rendered unimportant by the use of the large time interval.

cable, the average and most probable values of the measured momentum are to be identified with each other and with the local momentum at the momentary position of the center of the packet. This follows from the work of Sec. 12. In Sec. 33b, it will be proved that the statistical mean value of the square of the measured momentum for any single energy wave function is equal to the mean value of the square of the classical local momentum.

In connection with b the question arises: How are we to determine the wave function of an assemblage and to what extent can this be done experimentally? To answer this question we note that to prepare an assemblage of systems in a definite subjective state we must ordinarily start with a natural assemblage whose members are distributed through many such states and subject it to a sorting process which will eliminate all systems except those in a single, or very nearly single, subjective state (cf. Sec. 41d). In such cases the method of preparing the assemblage experimentally defines the state and we can often determine the wave function which characterizes it by elementary theory. Thus any classical scheme for preparing a beam of electrons with a definite energy is a scheme for preparing an assemblage whose wave functions all have a single vibration frequency. A beam of electrons diverging from a point source O will be associated with wave functions in the form of diverging spherical waves. If we block off all electrons from such a beam except those in the neighborhood of a point P which is far from O, the remainder of the assemblage will consist of electrons in states with Ψ functions in the form of wave packets with their normals approximately parallel to $\overrightarrow{OP}$.

Suppose, for example, that we wish to perform a simple electron-diffraction experiment. The electrons can be accelerated through a known potential difference V starting from an initial state in which their thermal energy is small compared to V. After passing through an initial aperture A they can be projected against a diaphragm B containing an aperture, or apertures, S. Near S the Ψ function for the electrons incident on B will have the form of a plane monochromatic wave. To determine the exact form of the wave function for the electrons emerging from S it would be desirable to treat quantum-mechanically the problem of the interaction of the incident wave system and the diaphragm B. In practice, however, it would probably be sufficient to adopt the procedure customary in the elementary theory of optical diffraction, assuming that for the emerging electrons, in the plane of B, Ψ vanishes on the diaphragm and has sensibly the same value at each point of the aperture as if the diaphragm were absent. The process of picking out of a natural assemblage of electrons a subassemblage with a definite wave function is essentially the same as that of preparing a beam or packet of *coherent*

radiation. The phase of the wave function is never determined, and in fact we may always regard the assemblage as an aggregate of sub-assemblages with Ψ's which differ in phase but are otherwise the same.

The above discussion assumes that the wave function of an assemblage of particles in a definite subjective state is to be determined theoretically from the method of preparing the state. It is also possible in principle to determine $|\Psi(x,t)|^2$ for such an assemblage by repeated positional observations on systems which belong to it, using high-frequency radiation.[1] Of course each such observation would alter the energy and momentum of the system observed and so exclude it from the original assemblage. If all members of the original assemblage were in the same state, however, the removal of a random selection of systems from the assemblage would not alter that state or the wave function which describes it.[2]

It follows that, for a pure case assemblage, $|\Psi(x,t)|^2$ is an essentially observable function and we infer that $\partial|\Psi(x,t)|^2/\partial t$ is also an observable function. Now Feenberg[3] has shown that the values of $|\Psi|^2$ and $\partial|\Psi|^2/\partial t$ at any given time t determine Ψ itself apart from a trivial constant phase factor. Thus, setting aside the physically meaningless phase factor, we can say that, in principle, $\Psi(x,t)$ can be determined by positional observations.

Feenberg's argument for the three-dimensional case—with slight modification—is as follows. Let

$$\Psi(x,y,z,t) = \chi(x,y,z,t)e^{i\varphi(x,y,z,t)}$$

where χ and φ are assumed to be real and analytic. Then [cf. Eq. (8·6)],

$$\frac{\partial}{\partial t}(\Psi\Psi^*) = \frac{h}{4\pi\mu i} \operatorname{div} [\Psi^* \operatorname{grad} \Psi - \Psi \operatorname{grad} \Psi^*],$$

or

$$\frac{\partial\chi^2}{\partial t} = \frac{h}{2\pi\mu} \operatorname{div} [\chi^2 \operatorname{grad} \varphi]. \qquad (15\cdot26)$$

Let χ^2 and $\partial\chi^2/\partial t$ be given for some time, say t_0. Then the function φ for t_0 is determined to an additive constant. To prove this we let u denote the difference between two possible solutions φ_1, φ_2 of the equation

$$\left(\frac{\partial\chi^2}{\partial t}\right)_{t_0} = \frac{h}{2\pi\mu} \operatorname{div} [\chi(t_0)^2 \operatorname{grad} \varphi].$$

We have

$$\operatorname{div} [\chi(t_0)^2 \operatorname{grad} u] = 0,$$

and hence

$$\operatorname{div} (u\chi^2 \operatorname{grad} u) = \chi^2 \operatorname{grad}^2 u.$$

---

[1] In virtue of the Pauli exclusion principle this statement holds even when the system contains two or more identical particles (cf. Sec. 42b).

[2] Cf. J. VON NEUMANN, *M.G.Q.*, p. 159.

[3] E. FEENBERG, "The Scattering of Slow Electrons in Neutral Atoms," *Thesis*, Harvard University, 1933.

Then, applying the transformation of Gauss, and observing that the resulting surface integral approaches zero as the volume over which we integrate approaches all configuration space as a limit, we obtain

$$\iiint_\alpha \chi^2 \, \mathrm{grad}^2 u \, d\tau = 0.$$

As the integrand is positive and continuous, it follows that χ grad u must vanish everywhere. But χ can vanish only on the nodal surfaces of Ψ; hence u must be a constant, say α.

Let χ_1 and χ_2 denote two possible choices of the function $\chi(x,y,z,t_0)$, consistent with a given form for $\chi(t_0)^2$. As these functions are real and analytic, it is necessary that

$$|\chi_2| = |\chi_1| \qquad \text{or} \qquad \chi_2 = \chi_1 e^{i\pi}.$$

In either case the complete wave function $\Psi(t_0) = \chi(t_0)e^{i\varphi(t_0)}$ is seen to be completely determined by $\chi^2(t_0)$ and $\left(\dfrac{\partial \chi^2}{\partial t}\right)_{t_0}$ except for a constant phase factor.

We conclude that the wave function of an assemblage known to be in a pure state is essentially observable.

16. THE HEISENBERG UNCERTAINTY PRINCIPLE[1]

The preceding paragraphs have made it clear that normally an assemblage described by a wave function Ψ involves a range of positional coordinates and momenta. It is not possible for a quadratically integrable wave function to represent an assemblage of systems having a unique value of the momentum, for single momentum functions are infinite plane waves. Thus a reduction in the range of momentum values associated with a wave packet tends to increase the range of coordinate values and *vice versa*. Heisenberg has given this principle the following quantitative formulation:

Let Δq_k and Δp_k denote respectively the uncertainties in the values of the coordinates q_k and of the corresponding momenta p_k computed from a wave function Ψ. The value of the product $\Delta q_k \Delta p_k$ varies with the choice of Ψ, but has the lower bound $h/4\pi$.

Here we define Δq_k and Δp_k as the respective root mean squares of the deviations of q_k and p_k from their mean values. To be precise, let Φ denote the probability amplitude for momentum defined in the three-dimensional case by Eq. (15·10) and let $d\tau_p$, $d\tau_q$ denote elements of volume in momentum space and the space of the coordinates q, respectively. Let the mean values of q_k and p_k be given by

$$\overline{q_k} = \int_\infty q_k |\Psi|^2 d\tau_q, \qquad \overline{p_k} = \int_\infty p_k |\Phi|^2 d\tau_p.$$

Then Δq_k and Δp_k are given by the relations

$$(\Delta q_k)^2 = \int_\infty (q_k - \overline{q_k})^2 |\Psi|^2 d\tau_q, \qquad (\Delta p_k)^2 = \int_\infty (p_k - \overline{p_k})^2 |\Phi|^2 d\tau_p. \quad (16\cdot1)$$

[1] *Cf.* W. HEISENBERG, *Zeits. f. Physik* **43**, 172 (1927); N. BOHR, *Nature* **121**, 580 (1928).

The relation $\Delta p_k \Delta q_k \geqslant h/4\pi$ is conveniently referred to as *Heisenberg's inequality*. A general proof of this inequality insofar as it applies to linear momentum is given in Sec. 33. Here we content ourselves with showing that in the case of a special type of one-dimensional wave packet previously discussed in Sec. 9, in which the distribution functions $|\psi(q,t)|^2$ and $|\Phi(p,t)|^2$ are Gaussian error curves, the product $\Delta q \Delta p$ has $h/4\pi$ as its minimum value. We identify q and p/h with the x and σ of Sec. 9, respectively. In that article it was proved that if Ψ has the form

$$\Psi(x,t) = \int_{-\infty}^{+\infty} G(\sigma)e^{2\pi i(x\sigma - \nu t)}d\sigma$$

appropriate to a free particle, and if $G(\sigma)$ is chosen to give all the component waves which form the integral the same phase at some space-time point x_0, t_0, the quantity Δx defined by (9·15) in harmony with (16·1) takes on a minimum value at $t = t_0$. We further proved that if $G(\sigma)$ is given the form of Eqs. (9·14) and (9·21), so that the momentum distribution function $|\Phi(p)|^2 = h^{-3}|G(\sigma)|^2$ has the form of a Gaussian error curve, the minimum value of Δx is $\dfrac{1}{\sqrt{2}}\dfrac{\alpha}{2\pi}$ [*cf.* Eq. (9·22)]. But Eq. (9·20) shows that under the same circumstances

$$(\Delta x)_{\min}{}^2 = \frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\left(\frac{dF}{d\sigma}\right)^2 d\sigma = \frac{\alpha^4}{4\pi^2}\int_{-\infty}^{+\infty}(\sigma - \bar\sigma)^2 F'^2 d\sigma = \frac{\alpha^4\overline{(\sigma - \bar\sigma)^2}}{4\pi^2}. \quad (16\cdot2)$$

Identifying $\overline{(\sigma - \bar\sigma)^2}$ with $(\Delta p/h)^2$, and combining (9·22) with (16·2), we obtain $(\Delta x)_{\min}\Delta p = h/4\pi$, independent of the choice of the parameters α, x_0, t_0.

Although the assumption that Ψ is a solution of the free-particle equation (9·1) was needed in Sec. 9 for a discussion of the motion of the packet and of the variation in Δx with time, the reader will readily verify that $\Delta x \Delta p = h/4\pi$ holds true at the instant t_0, whether the particle is free or not, provided that at the instant under consideration

$$\Psi(x,t_0) = \pi^{-\frac{1}{4}}\alpha^{\frac{1}{2}}e^{i\varphi_0}\int_{-\infty}^{+\infty} e^{-\frac{\alpha^2(\sigma - \sigma_0{}^2)}{2}+2\pi i(x - x_0)\sigma}d\sigma.$$

The relation $\Delta q \Delta p \geqslant h/4\pi$ is also valid if we identify q with the time and p with the energy of the system. Thus the uncertainty in the time at which a particle associated with a wave packet moving along the x axis passes the point $x = x_1$ would be approximately equal to the uncertainty in x divided by the group velocity $\partial\nu/\partial\sigma_x$.

$$\Delta t \cong \Delta x \div \frac{\partial\nu}{\partial\sigma_x} = \Delta x \div \frac{\partial E}{\partial p_x}.$$

It follows from the energy formula for a free particle, *viz.*,

$$E = \frac{p_x{}^2 + p_y{}^2 + p_z{}^2}{2\mu},$$

that, when the particle is traveling along the x axis (average values of p_y and p_z being zero), the uncertainty in E is produced almost entirely by the uncertainty in p_x. Then

$$\Delta E = \Delta p_x \frac{\partial E}{\partial p_x}$$

and

$$\Delta E \Delta t = \Delta x \Delta p_x \geqslant \frac{h}{4\pi}. \tag{16.3}$$

This result becomes reasonable when we recollect that in the Hamiltonian theory of classical dynamics t and $-E$ are canonically conjugate variables like x and p_x (*cf.* Sec. 7c). A plausible extrapolation would lead us to suppose that Heisenberg's uncertainty principle is applicable to any pair of classically conjugate dynamical variables. Robertson[1] has actually extended the proof to the general case where the conjugate variables p and q are functions of the Cartesian coordinates and momenta expressible as power series in the latter. Some difficulty is to be anticipated in the case of still broader generalizations, however, for not every pair of classically conjugate variables leads to a satisfactory corresponding pair of conjugate quantum-mechanical variables (*cf.* Sec. 39a, pp. 295–298).

We turn our attention now from the mathematical formulation of Heisenberg's uncertainty principle as a consequence of the hypotheses of wave mechanics to the experimental implications of the principle.

Since, by hypothesis, our maximum knowledge of the state of a system is given by a suitable wave function Ψ, it will be evident that, if the theory is correct and observations are made on a particle to determine the simultaneous values of a coordinate q and momentum p, the experimental uncertainties in the observed values must also be subject to the limiting relation

$$(\Delta q \Delta p)_{\text{min}} \sim h \tag{16.4}$$

The question now arises: Does this theoretical limitation on the accuracy of simultaneous observations of conjugate variables correspond to the experimental facts, or is it a weakness of the theory? This question has been answered by Heisenberg, who has shown by an analysis of the various possible experimental means for determining the values of coordinates and momenta that the former alternative is undoubtedly correct. For example, one can determine the position of a particle by observing the direction of motion of photons or electrons which have been scattered by collision with it, or one may allow the particle to pass through a small aperture in order to locate a point in its orbit. The momentum can be determined by successive positional observations in a field-free

[1] H. P. Robertson, *Phys. Rev.* **34**, 163 (1929).

space or by the Doppler effect. In each case if the alteration in momentum due to collision with a second particle is taken into account (*cf.* p. 59), and if the ordinary theory of optical diffraction is assumed to apply to matter corpuscles as well as to photons, it turns out that the relation (16·4) is verified by the analysis. For the detailed study of the various hypothetical and idealized experiments used in reaching this conclusion, the reader is referred to the original papers of Heisenberg and Bohr[1] and to the excellent exposition in Heisenberg's book *The Physical Principles of the Quantum Theory.*

Of course the analysis of Heisenberg's mental experiments is based on just those empirical facts regarding the Compton effect and the diffraction of matter which the wave mechanics is designed to describe. Hence the two methods of deriving the relation (16·4) are not altogether independent, but it is a source of satisfaction to find that the uncertainty principle can be established by very elementary considerations as well as from the detailed mathematical structure we have built up.

To avoid misunderstanding, a few words are introduced here about the meaning of the concept "observation." A more complete treatment of the subject will be given in Chap. IX.

When an atomic system is "observed" it interacts with an observing mechanism. If a certain type of observation is carried out for each member of an assemblage of similar systems, the assemblage will be divided, in general, into two or more subassemblages according to the outcome of the individual measurements. Thus, in the simple case of a beam of electrons radiating from an initial aperture A toward a diaphragm containing a second aperture S, the diaphragm is the essential feature of the observing mechanism and divides the electrons into two classes, *viz.*, those that get through and those that do not. We discover electrons of the former class beyond the diaphragm and say that they "have been observed to pass through S." Obviously the Ψ function which characterizes the future behavior of the subassemblage of electrons which have passed through the aperture is different from that which characterized the behavior of the complete assemblage leaving the initial aperture A. The discontinuous change in the wave function describing the subjective state of a system, which occurs when the system is measured, is called the "reduction of the wave packet" and will be discussed more fully in Sec.41*d*. The relation (16·4) applies to the values of $(\Delta q \Delta p)_{min}$ associated with either wave function, but not to mixed products involving, say, the Δp value for an assemblage before measurement and the Δq value for a subassemblage after measurement. Such mixed products can be made as small as desired but are of no particular physical significance.

Although a wave function Ψ defining the subjective state of a system does not *in general* permit a unique prediction regarding the result of any

[1] See footnote 1, p. 72.

particular observation or measurement of the system under considera-
tion,[1] it does give, as we have repeatedly stated, the maximum obtainable
information regarding the system. From the standpoint of classical
mechanics the state of a system is defined by the instantaneous values
of the coordinates and their corresponding momenta. Heisenberg's
uncertainty principle shows that actually this "classical state" can
never be known and hence has no operational meaning. The conception
of a particle with a uniquely defined position and momentum is a very
natural and useful extrapolation or idealization of everyday experience,
but in the last analysis we are driven by the experimental facts to the
conclusion that it has no place in the theory of atomic physics.

As Bridgman[2] has pointed out, our inability to make simultaneous
exact observations of position and momentum may be referred to our
inability to trace out the details of collisions between photons, electrons,
and apertures, or other more complicated systems. This inability,
in turn, is due to the absence of tools finer than complete collisions for
making the measurements necessary to give reality to such details.
The importance of this observation lies in the light which it sheds on the
nature of the new experimental discoveries which would be needed to
restore the validity of the classical conceptions of position and velocity.

It should be emphasized, however, that since simultaneous exact
values of positional coordinates and their conjugate momenta are
actually unthinkable in terms of wave-mechanical concepts, the discovery
of experimental means for breaking down the Heisenberg uncertainty
principle would not lead to an elaboration of our present quantum
mechanics but to the complete destruction of its essential features.
Hence we may regard the experimental evidence for the validity of
quantum-mechanical predictions as evidence that no such violations
of the uncertainty principle will ever be found.

The universal interest which the uncertainty principle has aroused
among physicists and philosophers is due to its intimate connection
with the hypothesis of indeterminism. So far as physics is concerned,
this hypothesis may be said to have originated in the discovery of the
law of radioactive decay. It received important support from Einstein's
speculations on the transition probabilities which govern the jumps of
atoms from one energy level to another and is the cornerstone of the
present theory (*cf.* Sec. 2). Classical determinism is the doctrine that,
if the present state of an isolated system (eventually the universe)
were known, the future behavior of the system could be exactly foretold
by a sufficiently expert mathematician. This view is incompatible with

---

[1] The same information is contained in the probability amplitude $\Phi(p_x, p_y, p_z, t)$
of Sec. 15 and in any of the infinite number of probability amplitudes obtainable from
Ψ by an appropriate change of variables.

[2] P. W. Bridgman, *Harper's Mag.*, March, 1929, p. 443.

quantum mechanics, which not only denies the validity of the hypothetical calculation but through the principle of uncertainty asserts that the information required by the supermathematician is both unobtainable and meaningless in terms of operations which we can perform, even in principle.

Although quantum mechanics is thus essentially indeterministic, a modicum of determinism remains in the fact that the Ψ function of an isolated system is uniquely determined for all time by its initial form (*cf.* Sec. 5). Indeterminism comes into the scheme in the relation between the wave function and the results of individual experiment, or measurement. As measurements always imply interaction of the system under observation with an external apparatus, it is possible to argue that the root of the indeterministic form of quantum mechanics lies in the apparently inevitable division of the universe into "observed" and "observer."

CHAPTER III

ONE-DIMENSIONAL ENERGY-LEVEL PROBLEMS

17. BOUNDARY AND CONTINUITY CONDITIONS; EIGENVALUES AND EIGENFUNCTIONS

It is a well-known empirical fact that each species of atom has a characteristic set of discrete energy levels and associated "stationary" states. The energy differences of these states determine the frequencies of the spectrum lines and the so-called *transition probabilities* for "jumps" between the levels determine the intensities. Perhaps the most fundamental task of quantum mechanics is that of locating these levels and evaluating the transition probabilities on a purely theoretical basis.

Since we have already identified energy and frequency, it is clear that the state of an atom having a definite energy must be given by a *monochromatic* solution of the corresponding Schrödinger equation. If the *internal* energy is to be definite, the wave function for the atom must factor into the product of a function of the coordinates of the center of gravity, Ψ_g, and a function of the relative coordinates, Ψ_r [*cf.* Eq. (15·16)]. Then Ψ_r must be a monochromatic solution of Eq. (15·18) and its amplitude function ψ_r must be a solution of the time-free wave equation

$$\sum_{k=2}^{n} \frac{1}{\mu_k}\left(\frac{\partial^2}{\partial \xi_k{}^2} + \frac{\partial^2}{\partial \eta_k{}^2} + \frac{\partial^2}{\partial \zeta_k{}^2}\right)\psi_r + \frac{1}{\mu_1}\sum_{j=2}^{k}\sum_{k=2}^{n}\left(\frac{\partial^2}{\partial \xi_j \partial \xi_k} + \frac{\partial^2}{\partial \eta_j \partial \eta_k} + \frac{\partial^2}{\partial \zeta_j \partial \zeta_k}\right)\psi_r$$
$$+ \frac{8\pi^2}{h^2}(E - V)\psi_r = 0. \quad (17 \cdot 1)$$

When this equation is applied to an atom with its single heavy nucleus and group of $f = n - 1$ identical electrons, it is convenient to identify the nucleus with particle 1. The coordinates ξ_k, η_k, ζ_k are then the Cartesian coordinates of the electrons referred to a set of axes with the origin at the nucleus. Usually the situation is idealized by treating the nuclear mass as infinite in comparison with the electronic mass μ.[1] Then the above equation reduces to

$$\sum_{k}\left(\frac{\partial^2}{\partial \xi_k{}^2} + \frac{\partial^2}{\partial \eta_k{}^2} + \frac{\partial^2}{\partial \zeta_k{}^2}\right)\psi_r + \frac{8\pi^2\mu}{h^2}(E - V)\psi_r = 0 \quad (17 \cdot 2)$$

Here the summation is to be extended over all the electrons; they may be renumbered so that k runs from 1 to f.

[1] *Cf.*, however, the correction for the finite mass of the nucleus given by D. S. Hughes and C. Eckart, *Phys. Rev.* **36**, 694 (1930).

In order to get discrete energy levels from Eq. (17·1) we must seek a class of solutions which satisfy appropriately chosen *boundary conditions*. Otherwise all energy values would stand on the same footing. The boundaries to be considered are the "point" at infinity and the points or domains where the potential energy becomes infinite. Such domains are called *singular domains*[1] of the differential equation and include all points of configuration space which locate any two particles at the same point of ordinary three-dimensional space. There is a marked tendency for solutions of the differential equation to become infinite at the boundary points, and therefore Schrödinger introduced as a boundary condition the requirement that Ψ shall be finite everywhere. This condition usually gives energy values in conformity with experimental results, but does not apply to the Dirac relativistic theory of the single electron, and is too much of an *ad hoc* proposition to be wholly satisfactory in any case.

A better motivated condition is that of *quadratic integrability*, by which we mean the requirement that $\int \Psi \Psi^* d\tau$ when extended over all configuration space shall converge. This condition is forced upon us, as we have already seen (*cf.* Sec. 8), by the fundamental hypothesis that $\Psi \Psi^* d\tau$ measures the probability that the system represented by Ψ has a configuration belonging to the element $d\tau$ of configuration space. In the neighborhood of the point at infinity the quadratic integrability requirement is stronger than the Schrödinger boundary condition, but at finite singular points the Schrödinger condition is stronger.

For the present it will be convenient to make use of both of these restrictions and of the following continuity condition.

Continuity Condition: All wave functions are required to be continuous and single-valued, and to possess continuous first and second derivatives with respect to the Cartesian coordinates in the neighborhood of every interior point of the region of definition of the Schrödinger equation (i.e., in the neighborhood of every nonsingular point).

Functions which satisfy the continuity conditions and are bounded and quadratically integrable will be referred to as functions of *class* or *type A*. Such functions will also be referred to as "physically admissible," by which we indicate that they are adapted to the description of the

---

[1] The existence theorem for an ordinary linear differential equation of the second order such as Eq. (23·4) below states that if the coefficients p_0, p_1, p_2, ρ are real and continuous in the neighborhood of a point $x = a$, and if p_0 does not vanish at that point, there is a solution continuous in the neighborhood of $x = a$ and satisfying the arbitrary initial conditions

$$y(a) = y_0, \qquad y'(a) = y_0'.$$

A *singular point* for such a differential equation is defined as a point at which one of the conditions for the establishment of the existence theorem ceases to hold. (*Cf.* E. L. Ince, *Ordinary Differential Equations*, pp. 69, 72, 160, 356–370. London. 1927.)

behavior of quantum-mechanical physical systems. In Sec. 32*b* the question of the exact nature of the boundary-continuity conditions appropriate to physically admissible functions will be reconsidered and new and more stringent requirements will be laid down.

In addition to the type *A* functions we shall have use for solutions of the Schrödinger equations which are not quadratically integrable and hence not physically admissible in the above sense, but which do satisfy the continuity condition and can be compounded by integration to build up wave-packet solutions of the Schrödinger equation (7·3) which *are* physically admissible. We shall designate functions of this latter variety as functions of *class* or *type B*. The plane-wave functions of Chap. II are simple examples.

The calculation of the energy levels is thus reduced to a boundary-value problem of a type long familiar to mathematical physicists. (The simplest example of a classical problem of this type is the determination of the frequencies and modes of vibration of a stretched string with fixed end points.) Type *A* solutions of Eqs. (17·1) and (17·2) exist only for certain discrete values of *E*. In addition there are type *B* solutions for a continuous range of *E* values, which usually extends from a finite lower limit to infinity. All of the energy values which meet the milder condition *B* are called by various authors "characteristic values," "proper values," or "eigenvalues"[1] and can be identified with energy levels of the atom (or molecule) under consideration. The corresponding solutions of the differential equation are called "characteristic functions," "proper functions," or "eigenfunctions." The problem of determining a set of eigenvalues and eigenfunctions from a differential equation is closely paralleled by certain problems in the solution of simultaneous linear algebraic equations involving a parameter. Hence we speak of "eigenvalue problems" in algebra as well as in differential equations. Thus the determination of the lengths and directions of the principal axes of a general ellipsoid is an algebraic eigenvalue problem while the determination of the frequencies of vibration of an elastic system is a familiar type of eigenvalue problem based on the differential equations of classical physics. Still another type of eigenvalue problem arises in the domain of integral equations, into which differential equations can often be converted.

In this chapter we make contact with a highly developed field of classical mathematics and mathematical physics. The author's endeavor is to sketch in as brief a manner as possible those elements of this older, but still developing, mathematical structure which are essential to

---

[1] Half-translated from the German "*Eigenwerte.*" We shall use the terms "eigenvalue" and "eigenfunction" in conformity with the custom of many other English-speaking physicists although we regard the adoption of such a hybrid nomenclature as unfortunate.

quantum mechanics. Readers to whom eigenvalue problems are unfamiliar are urged to supplement the account here given with extensive reading in the mathematical texts referred to in footnote 1, p. 90. More advanced students are referred to the highly abstract but systematic general attack on the eigenvalue problems of the quantum theory (the problem of the parametric Schrödinger equation is not the only one!) in von Neumann's book *Mathematische Grundlagen der Quantenmechanik* (*cf.* book list, p. xviii).

18. THE ONE-DIMENSIONAL ANHARMONIC OSCILLATOR

As a first example of the energy-level problem we consider the special case of a particle of mass μ vibrating in one dimension under the influence of forces derivable from a potential-energy function $V(x)$. We assume that $V(x)$ has a single minimum placed at the origin for convenience, a pole at $x = -\infty$, and a finite asymptotic limit D at $x = +\infty$ (*cf.*
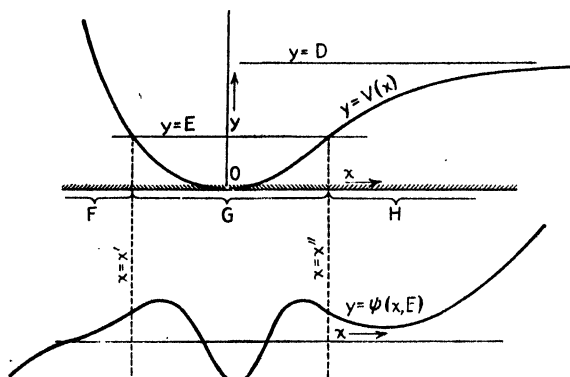
FIG. 4.—Potential energy and integral curve graphs for anharmonic oscillator.

Fig. 4). In discussing this problem it is convenient to consider three cases corresponding to different ranges of energy values, as follows:

Case I: $E < V_{\min} = 0$; Case II: $0 < E < D$; Case III: $D < E$.

Case I is classically impossible. As we shall see, it yields no solutions which conform to the boundary conditions. Case II yields oscillatory motions in classical theory and gives a series of discrete quadratically integrable eigenfunctions in quantum mechanics. In Case III the classical motion is aperiodic so that if one knew the energy of the particle and nothing more, the expectation of finding it in any finite interval of the x axis would be nil. The eigenfunctions for Case III are of the type B, not quadratically integrable, from which we again infer that the expectation of finding a particle of sharply defined energy greater than D in any finite interval of the x axis is zero, being the ratio of the integral of $\psi\psi^*$ over the interval in question to the integral of the same quantity from $-\infty$ to $+\infty$.

The classical local momentum and corresponding local de Broglie wave length are given by the formulas

$$p(x,E) = \frac{h}{\lambda(x,E)} = \sqrt{2\mu[E - V(x)]}. \tag{18·1}$$

The corresponding wave equation (5·9) for the space factor ψ reduces to the form

$$\frac{d^2\psi}{dx^2} + \kappa(E - V)\psi = 0. \qquad \kappa = \frac{8\pi^2\mu}{h^2} \tag{18·2}$$

λ has a minimum value at $x = 0$ and becomes infinite at points where $V(x)$ is equal to E (e.g., the points x' and x'' in Fig. 4). Outside the region G in which $V(x)$ is less than E the formulas yield imaginary values of the wave length and momentum. Moreover, if we compute the kinetic energy by the plausible formula

$$T = \frac{p^2}{2\mu} = E - V = -\frac{1}{\kappa\psi}\frac{d^2\psi}{dx^2},$$

we obtain a negative value.

Such regions of *imaginary classical local momentum* are excluded in classical mechanics but play an important part in the present theory. The wave functions of our problem are spread out over the whole of the x axis including the part outside the region G. It follows from the physical interpretation of $|\psi|^2$ that if the particle is in a state of definite energy E, there is a positive probability that a measurement of its position will discover it in some part of the region of imaginary local momentum. This does *not* mean, however, that measured momentum as defined in Sec. 15 can be imaginary, or that the measured kinetic energy, as computed from the measured momenta, can be negative. The relative probabilities of different measured values of these quantities are *not* point functions of the coordinates but properties of the complete wave function of the system.

19. THE QUALITATIVE BEHAVIOR OF THE INTEGRAL CURVES:[1] EXISTENCE OF CLASS A EIGENFUNCTIONS

19a. Behavior of Integral Curves in Regions of Positive and Negative Kinetic Energy.—As the coefficients of Eq. (18·2) are real, the real and imaginary parts of every eigenfunction must be eigenfunctions when taken separately. Any two type A or type B solutions of the equation for a given energy value can be proved to be linearly dependent (*i.e.*, one is a multiple of the other), and hence any eigenfunction can be resolved into the product of a real function of x and a complex constant. *Consequently it suffices to search for real solutions of the problem* which have the advantage that they can be represented by single integral curves.

[1] *Cf.* F. HUND, *Zeits. f. Physik* **40**, 742 (1927).

Since the equation is of the second order, every solution is characterized by two constants of integration. It follows from a fundamental existence theorem that an integral curve representing a real solution of the differential equation can be passed through any point of the x, ψ plane with an arbitrary slope at that point. The ordinate and slope of the curve for the given abscissa may be identified with the constants of integration of the solution. In general, however, such an integral curve will not satisfy the boundary conditions imposed on eigenfunctions and we have to pick out from the complete manifold of integral curves for all values of E those which do conform to the boundary conditions.

Integral curves are obviously concave to the x axis where $\frac{1}{\psi} \frac{d^2\psi}{dx^2}$ is negative and convex to the axis where it is positive. Hence we can see from the differential equation that inside the region G, where E is greater than $V(x)$, the curves are oscillatory, crossing and recrossing the axis (*cf.* Fig. 4). In the regions F and H to the left and right of G, on the other hand, the curves are convex to the axis. Hence none of them can cross the axis more than once in either of these regions. The nodes of the integral curves and the points x' and x'' which separate G from F and H are points of inflection.

If $D - E$ is positive, so that the region H really exists, every solution of the differential equation must become infinite, or approach zero monotonically, as x increases from x'' to ∞. Similarly, every real solution of the differential equation either becomes infinite, or approaches zero monotonically, as x decreases from x' to $-\infty$. It follows that every type A eigenfunction must vanish at the "boundary points" $x = \pm \infty$. Conversely, every solution of the differential equation which vanishes at $x = \pm \infty$ is quadratically integrable and hence a type A eigenfunction. By an appropriate choice of constants of integration α and β we can always (independent of E) pick out an integral curve which meets the above boundary condition at one of the boundary points, say $x = -\infty$. Identifying α and β, respectively, with the ordinate and slope of the curve at some point $x = \xi$ in the region F, let us consider the family of integral curves obtained by holding α and E constant while varying β. For small positive values of β the curve $y = \psi(x)$ will first approach the axis as x takes on increasing negative values and then recede from it as x approaches $-\infty$. For large positive values of β the curve is sure to cross the axis and ψ becomes negatively infinite as x approaches $-\infty$. For a certain unique value of β, say β_1, there must be an intermediate curve which approaches zero monotonically as x approaches $-\infty$. There is then one, and only one, integral curve for each pair of values of α and E which vanishes at $x = -\infty$.[1] If we multiply the corresponding solution of the differential equation, say $y = \psi(x, E)$, by an arbitrary constant, we

---

[1] These statements are rigorously proved in Appendix C.

obtain a new solution with the same value of E which also satisfies the boundary condition at $-\infty$. We cannot in this way affect the behavior of the integral curves at the positive end of the x axis, however, and must look to a variation of the energy E for aid in picking out integral curves which meet the boundary conditions at both $+\infty$ and $-\infty$.

19b. The Discrete Eigenfunctions.—Consider next the way in which the adjustment of E affects the integral curves. Let the functions $\psi_1(x)$ and $\psi_2(x)$ denote integral curves for the energies E_1 and E_2, respectively, both of which satisfy the left-hand boundary condition. If $E_1 > E_2$, the curve $\psi_1(x)$ will oscillate more rapidly in G than $\psi_2(x)$ and its nodes will be closer together. Furthermore, it is easy to see that the first node of $\psi_1(x)$ (not counting the node at $-\infty$) lies to the left of the first node of $\psi_2(x)$. Hence each node of the former function will lie to the left of the corresponding node of the latter.[1] As E increases continuously the nodes move to the left, new nodes appear at $x = +\infty$ and enter the region G where they accumulate. The first appearance of each node at $x = +\infty$ marks a corresponding eigenvalue for E and eigenfunction ψ.

The minimum number of nodes between the boundaries is zero and the eigenfunction having no nodes except those at the boundary points has
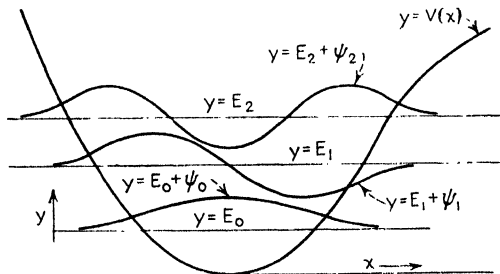
Fig. 5.—Eigenfunctions of the anharmonic oscillator.

the lowest eigenvalue of E. Since a continuous curve with nodes at $-\infty$ and $+\infty$ must be concave to the axis somewhere, the minimum eigenvalue must be greater than the minimum value of $V(x)$. Thus the eigenvalue-eigenfunction problem has no solutions for Case I.

Clearly any two eigenfunctions having the same eigenvalue must have the same zeros and can differ only by a multiplicative constant.[2] For our present purpose they can be identified with one another. If we number the eigenfunctions in the order of their energy values starting with zero, the nth of these functions will have exactly n internal nodes dividing the x axis into $n + 1$ parts. In this theory the number of

[1] For a more exact discussion of the displacement of the nodes as E increases see Riemann-Weber, *D.P.*, vol. I, p. 291.

[2] This is not usually true in multidimensional problems.

nodes plays a role analogous to that of the integral quantum numbers of the Bohr theory. It is an integer which is used to identify the eigenvalues and eigenfunctions, giving the ordinal number of the former when arranged in a series according to magnitude.

Figure 5 shows the qualitative form of the eigenfunctions for some of the lower eigenvalues.

19c. The Continuous Spectrum of Class B Eigenfunctions.—The above discussion covers the energy range of Case II. In Case III, where E is greater than D, the integral curves cannot be made to vanish at $x = \infty$. For large values of x we have approximately (*cf.* p. 87).

$$\frac{d^2\psi}{dx^2} = -\alpha^2\psi. \qquad \alpha^2 = \kappa(E - D) > 0.$$

Hence ψ has the asymptotic form[1]

$$\psi = C \cos (\alpha x - \epsilon),$$

appropriate to a free particle, and we conclude that it cannot be made quadratically integrable for any value of E greater than D. Thus the spectrum of discrete energy levels whose eigenfunctions satisfy the boundary conditions A terminates at $E = D$. However, the asymptotic form $C \cos [\alpha(E)x - \epsilon(E)]$ shows that all solutions of the wave equation for positive values of $E - D$ remain finite as x approaches ∞. By adjustment of the constants of integration we can obtain a solution for every energy which vanishes at $x = -\infty$ and has any desired amplitude $C(E)$ at $x = +\infty$. If $C(E)$ is made a continuous function, the integral curves derived in this way form a one-parameter family $\psi(x,E)$ which depends continuously on E. The integration of this family over a range of energy values gives a function $\Delta\psi$ which is readily shown to vanish at $x = +\infty$ as $1/x$, and which is therefore quadratically integrable. Thus, using the asymptotic form we have for large values of x,

$$\Delta\psi = \int_{E_1}^{E_2} C(E) \cos [\alpha x - \epsilon]dE = \int_{\alpha_1}^{\alpha_2} A(\alpha) \cos \alpha x \, d\alpha + \int_{\alpha_1}^{\alpha_2} B(\alpha) \sin \alpha x \, d\alpha,$$

where

$$A(\alpha) = U(E) = C(E) \cos \epsilon \frac{dE}{d\alpha}; \qquad B(\alpha) = V(E) = C(E) \sin \epsilon \frac{dE}{d\alpha}.$$

Successive integrations by parts yield an expansion in inverse powers of x:

$$\Delta\psi = \frac{1}{x}[A(\alpha) \sin \alpha x - B(\alpha) \cos \alpha x]\Big|_{\alpha_1}^{\alpha_2} + \frac{1}{x^2}\left[\frac{dA}{d\alpha} \cos \alpha x + \frac{dB}{d\alpha} \sin \alpha x\right]\Big|_{\alpha_1}^{\alpha_2} + \cdots.$$

This expansion shows that $\Delta\psi$ is quadratically integrable. *It follows*

[1] For a rigorous discussion of the asymptotic form of $\psi(x)$ see Theorem II, Appendix H.

that $\psi(x,E)$ is a class B function. (By making $C(E)$ a differentiable function which vanishes at E_1 and E_2, we can eliminate the terms in $1/x$ in the above series. In fact we can apparently make $\Delta\psi$ approach zero at infinity as rapidly as any desired power of $1/x$ by choosing $C(E)$ as an analytic function which vanishes together with its first n derivatives at E_1 and E_2. This conclusion, however, takes inadequate account of the discrepancies between $\psi(x,E)$ and the corresponding asymptotic B. W. K. approximation.)

The continuous range of energy values for which $E > D$ will be referred to as the *continuous energy spectrum* and the corresponding type B solutions of the wave equation will be called *continuous spectrum*, or *type B, eigenfunctions.* This term will be applied to both the space factor ψ and the complete wave function $\Psi = \psi e^{-2\pi iEt/h}$.

If the potential function $V(x)$ is infinite at both extremes of the x axis the continuous energy spectrum disappears and the discrete spectrum has an infinite range of values. A special case of interest is that of the ideal linear oscillator where $V = \frac{1}{2}kx^2$. This problem is discussed in detail in Sec. 20. The same remark holds good if $V(x)$ has two poles of the second or higher order at finite points x_1,x_2, and if we consider solutions in the range $x_1 < x < x_2$. In this case ψ must vanish at the singular points or become infinite there. The forms of the lower eigenfunctions remain essentially the same independent of the behavior of $V(x)$ at points far from its minimum.

19d. The Paradox of the Nodes.—The existence of nodes in the eigenfunctions of the wave equation may be disconcerting at first sight, since it means that if we determine the positions of an assemblage of linear oscillators all of which have a common definite energy, we shall never find particles at the nodal points and very rarely near these points. If we are resigned to the undulatory theory of matter, the existence of these nodes is in itself no stranger than the existence of nodes in a standing wave system in the electromagnetic theory of light. But we are in apparent difficulty if we ask how it happens that we sometimes find a particle on one side of a node and sometimes on the other—never right at the node itself. For if we think of the particles as vibrating back and forth, we expect that unless they attain an infinite velocity in passing through the nodes, they must spend some time in the neighborhood of each.

. The paradox originates in an unwarranted though inevitable carryover of classical conceptions into the new mechanics. There is no way in which the motion of a particle having a definite energy can be followed experimentally, and hence we must infer that the classical conception of a definite space-time orbit is inapplicable. It is best to say that the particle has no definite position until its position is measured (*cf.* p. 76). When that is done the particle is left in a state whose wave function

forms a very compact wave packet and is no longer in the original energy level. We may perhaps think of position, momentum, and energy as properties which the particle assumes under the stress of appropriate experiments. Since it is not possible to measure position and energy exactly at the same time, we shall not assume the existence of either one unless Ψ has a special form which defines the quantity uniquely. This is about the same thing as saying that position and energy exist only when there is no uncertainty regarding their values.[1] In any case it should be clear that the paradox of the nodes involves no question of disagreement between theory and experiment but rather a disagreement between theory and metaphysical preconceptions. We return to these philosophical questions in Sec. 36b, and Chap. IX.

20. THE PLANCK IDEAL LINEAR OSCILLATOR

20a. The Sommerfeld Polynomial Method.—The simplest example of an energy-level problem whose exact solution has been worked out is that of the Planck ideal linear oscillator with the potential-energy function $V = \tfrac{1}{2}kx^2$. The corresponding wave equation and boundary-value problem were well known to mathematicians before the advent of quantum mechanics.[2]

To solve Eq. (18·2) for this special case we shall employ the polynomial method of Sommerfeld[2] which is based on the fact that in many cases each eigenfunction can be factored into the product of a polynomial $P(x)$ and a transcendental or irrational algebraic function $Q(x)$. The roots of the polynomial give the nodes of the function, and its degree is therefore equal to the number of nodes. The factor $Q(x)$ takes care of the behavior of the function at the singular points which bound the region in which the differential equation is defined.

For example, in the case under discussion the points $x = \pm \infty$, where V becomes infinite are singular points. To determine the appropriate form for $Q(x)$ we first seek an approximate solution of the differential equation valid only for large values of x^2. The equation is conveniently written in the form

$$\frac{d^2\psi}{dx^2} + (\beta - \alpha^2 x^2)\psi = 0 \qquad (20\cdot1)$$

[1] To say that the matter particles do not have definite positions except under these special conditions is really giving up our initial naive hypothesis that matter ultimately consists of corpuscles. It would hardly be legitimate to stretch the definition of a corpuscle to include an entity that had a definite position only a part of the time, and an infinitesimal part of the time at that. We substitute the postulate that matter is an entity which exhibits the characteristics of corpuscles whenever it is subjected to observations designed to measure its position.

[2] E. Schrödinger, *Ann. d. Physik* (4) **79**, 489 (1926); Courant-Hilbert, *M.M.P.*, p. 283; A. Sommerfeld, *Atombau und Spektrallinien, Wellenmechanischer Ergänzungsband*, p. 2, Braunschweig, 1929.

with the notation

$$\beta = \kappa E = \frac{8\pi^2\mu E}{h^2}; \qquad \alpha^2 = \frac{\kappa k}{2} = \frac{4\pi^2\mu k}{h^2}.$$

For large values of x^2 the term in β may be neglected and we can immediately write down the approximate solution

$$\psi = Q = e^{\pm\frac{\alpha x^2}{2}}.$$

In order to satisfy the boundary condition that $\psi(\pm\infty) = 0$, we must adopt the negative exponent. We next assume for the rigorous solution the form

$$\psi = ve^{-\frac{\alpha x^2}{2}}. \tag{20.2}$$

Substitution into Eq. (20·1) yields the differential equation

$$\frac{d^2v}{dx^2} - 2\alpha x\frac{dv}{dx} + (\beta - \alpha)v = 0 \tag{20.3}$$

for v. This equation can be solved by means of a power series in x. For convenience we first make the substitution

$$\xi = \sqrt{\alpha}x,$$

which yields

$$\frac{d^2v}{d\xi^2} - 2\xi\frac{dv}{d\xi} + \left(\frac{\beta}{\alpha} - 1\right)v = 0. \tag{20.4}$$

Introducing the series expansion $v = \sum_\tau a_\tau\xi^\tau$ and equating the coefficient of each power of ξ to zero we obtain the recurrence formula

$$(\tau + 2)(\tau + 1)a_{\tau+2} + \left(\frac{\beta}{\alpha} - 1 - 2\tau\right)a_\tau = 0.$$

This formula shows that a_0 determines the coefficients of all the even powers of ξ while a_1 determines the coefficients of all the odd powers. a_0 and a_1 themselves are arbitrary and may be regarded as constants of integration. If we set a_1 equal to zero we shall get a solution involving only even powers of x, and therefore as a whole an even function of x. If we set a_0 equal to zero, we get a solution which is an odd function of x. Since V is an even function of x and the boundary conditions are the same at the two ends of the x axis, it is evident that each of the eigenfunctions must be either an even or an odd function of x.[1] Hence we must expect that in every case either a_0 or a_1 will be zero.

[1] To give a rigorous proof, one may proceed as follows. Let $\psi = f(x)$ be a solution of Eq. (20·1) which satisfies the boundary conditions. Make the change of variables

$$x = -x'.$$

Then Eq. (20.1) becomes

20b. Determination of the Eigenvalues.—For very large values of τ we have approximately

$$\frac{a_{\tau+2}}{a_\tau} \cong \frac{2}{\tau + 2}.$$

This ratio occurs also in the series expansion of e^{ξ^2}, so that in general both the even and odd series become infinite as e^{ξ^2} when ξ becomes positively or negatively infinite. It follows that despite the exponential decrease of Q with increasing values of x^2 the complete function ψ usually becomes infinite as $e^{\alpha x^2/2}$ when x^2 becomes infinite. Exceptions occur for those special values of the energy which cause one series or the other to break off after a finite number of terms. Suppose, for example, that

$$\frac{\beta}{\alpha} = 2n + 1, \tag{20·5}$$

where n is any positive integer. Then the recurrence formula shows that a_{n+2} is zero. If n is even, the series breaks off with the term $a_n\xi^n$ and we get a solution by setting a_1 equal to zero. Similarly we get a solution for odd values of n by setting a_0 equal to zero. Inserting the values of α and β into Eq. (20·5) we obtain the eigenvalues of E,

$$E_n = (n + \tfrac{1}{2})h\nu_c, \tag{20·6}$$

where ν_c is the classical vibration frequency $(2\pi)^{-1}\sqrt{k/\mu}$. This is the same as the Bohr energy-level formula with half-integral quantum numbers.

20c. The Eigenfunctions and Their Properties.—Combining the recurrence formula with Eq. (20·5) we obtain the following explicit expressions for the coefficients in the polynomials $v_n(\xi)$:

$$\frac{a_\tau}{a_0} = \frac{(-2)^{\frac{\tau}{2}} n(n-2) \cdots (n-\tau+2)}{\tau !}, \qquad (n, \tau \text{ even})$$

$$\frac{a_\tau}{a_1} = \frac{(-2)^{\frac{\tau-1}{2}} (n-1)(n-3) \cdots (n-\tau+2)}{\tau !}. \qquad (n, \tau \text{ odd})$$

The v_n's so defined are identical to a constant factor with the well-known

---

$$\frac{d^2f}{dx'^2} + (\beta - \alpha^2 x'^2)f = 0.$$

This is identical in form with the original equation, showing that $f(-x)$ is, like $f(x)$, a solution of (20·1). As $f(-x)$ satisfies the boundary conditions, it is an eigenfunction with the same eigenvalue as $f(x)$. Hence it can differ from $f(x)$ at most by a constant factor. By the normalization condition this factor is ± 1. $f(x)$ is therefore either an even or an odd function of x.

Hermitian polynomials $H_n(\xi)$ which may also be written in the form

$$H_n(\xi) = (-1)^n e^{\xi^2}\frac{d^n e^{-\xi^2}}{d\xi^n}. \tag{20.7}$$

The eigenfunctions $\psi_n(\xi)$ are in turn identical with the Hermitian orthogonal functions. The properties of these functions are quite fully described in various mathematical texts.[1] The most important are the following:

$$H_{n+1} - 2\xi H_n + 2nH_{n-1} = 0; \tag{20.8}$$

$$\frac{dH_n(\xi)}{d\xi} = 2nH_{n-1}(\xi); \tag{20.9}$$

$$\int_{-\infty}^{+\infty} e^{-\xi^2}H_m(\xi)H_n(\xi)d\xi = 2^n n!\sqrt{\pi}\,\delta_{mn}. \tag{20.10}$$

In the last equation the Kronecker symbol δ_{mn} denotes that function of the integers m and n which is unity when $m = n$ and is zero when $m \neq n$.

Equation (20.10) is a statement of the quadratic integrability of the eigenfunctions and can be used to normalize them as required by Eq. (8.9). We write

$$\psi_n = c_n H_n(\xi)e^{-\frac{\xi^2}{2}},$$

where c_n is a complex constant containing a_0 or a_1, as the case may be. Then c_n is to have an absolute value such that

$$1 = \int_{-\infty}^{+\infty} \psi_n\psi_n{}^*dx = \frac{c_n c_n{}^*}{\sqrt{\alpha}}\int_{-\infty}^{+\infty} e^{-\xi^2}H_n{}^2(\xi)d\xi = \frac{c_n c_n{}^*}{\sqrt{\alpha}}2^n n!\sqrt{\pi}.$$

Solving for c_n and expressing α in terms of the vibrational frequency ν_c we obtain the expression

$$\psi_n(x) = e^{i\chi_n}\left(\frac{k}{\pi h\nu_c}\right)^{\frac14}\frac{1}{\sqrt{2^n n!}}H_n\left(x\sqrt{\frac{k}{h\nu_c}}\right)e^{-\frac{kx^2}{2h\nu_c}} \tag{20.11}$$

for the normalized ergenfunctions of the problem. χ_n is an arbitrary phase constant usually set equal to zero.

21. AN APPROXIMATION METHOD WHICH CORRELATES THE EIGENVALUES OF WAVE MECHANICS WITH THE ENERGY LEVELS OF THE BOHR THEORY

21a. The B. W. K. Approximations for $\psi(x)$.—The possibility of finding the experimental energy levels of atoms from wave theory was first indicated by de Broglie[2] who showed that the Bohr quantum condition

---

[1] COURANT-HILBERT, *M.M.P.*, pp. 77–79, 283; RIEMANN-WEBER, *D.P.*, vol. I, pp. 342–347.

[2] L. DE BROGLIE, *Thesis*, Chap. III; *J. de Physique* **7**, 327 (1926). *Cf.*, however, M. BRILLOUIN, *Comptes Rendus* **168**, 1318 (1919), **169**, 48 (1919), **171**, 1000 (1920); *J. de Physique* **3**, 65 (1922).

for circular orbits in hydrogen is identical with the condition that the "optical path length" around the orbit is an integral number of wave lengths. Later Schrödinger[1] showed that in a number of important special cases the energy values given by appropriate solutions of his first wave equation are in substantial agreement with the Bohr theory and with experiment. It remained for Brillouin, Wentzel, and Kramers,[2] however, to show why it is that the eigenvalue theory of wave mechanics is in such systematic agreement with the energy-level predictions of the Bohr quantum theory and to give an explanation of such discrepancies as do occur. They have developed a powerful method of attack on eigenvalue problems in which the Sommerfeld phase-integral formulas of the Bohr theory drop out in first approximation. The method is based on a simple idea and is relatively easy to apply, although a rigorous justification of the procedure (cf. Secs. 21c, d, e, and f) leads to rather tedious calculations. It is primarily adapted to problems which can be thrown into one-dimensional form by separation of the variables. Even when this cannot be done rigorously, it is frequently possible to introduce approximations which make such separation possible. We accordingly restrict the discussion here to the one-dimensional case.

Let us seek a fundamental pair of solutions of the Schrödinger equation (18·2) having the form[3]

$$\psi = A e^{\frac{2\pi i}{h} \int y(x, E) dx}, \tag{21·1}$$

where A is a constant. Direct substitution shows that y must satisfy the first-order Ricatti equation

$$\frac{h}{2\pi i} y' + y^2 = 2\mu(E - V), \tag{21·2}$$

in which y' denotes the derivative dy/dx. Conversely every solution of (21·2) yields a corresponding solution of the wave equation (18·2).

In order to solve (21·2) Brillouin and Wentzel independently suggested a development in power series in h. As quantum mechanics passes over into classical mechanics when h is set equal to zero, it is to be expected that early approximations of this type will prove most useful in the realm of high quantum numbers where classical theory is most nearly correct. Unfortunately an expansion of this kind is only semiconvergent and cannot yield an exact solution of the problem.[4] On this account we

[1] E. Schrödinger, *Ann. d. Physik* (4) **79**, 361, 489 (1926).

[2] L. Brillouin, *Comptes Rendus* **183**, 24 (1926); *J. de Physique* **7**, 353 (1926); G. Wentzel, *Zeits. f. Physik* **38**, 518 (1926); H. A. Kramers, *Zeits. f. Physik* **39**, 828 (1926). The mathematical method of Brillouin, Wentzel, and Kramers seems to have been first used by H. Jeffreys, *Proc. London Math. Soc.* (2) **23**, 428 (1923).

[3] *Cf.* the substitution (11·2) leading to the Hamilton-Jacobi equation (11·4).

[4] *Cf.* A. Zwaan, *Arch. Néerland.*, ser. IIIA, tome XII, 1 (1929); G. D. Birkhoff, *Bull. Am. Math. Soc.* **39**, 696 (1933); R. E. Langer, *ibid.* **40**, 545 (1934).

shall limit the discussion here to the approximation obtained by using the first two terms of an expansion of y in powers of h. What we lose thereby in generality will be offset by the fact that we can write down an explicit formula for the approximation under consideration in terms of $V(x)$.[1]

Inserting the series $y = \sum_{k=0}^{\infty} \left(\dfrac{h}{2\pi i}\right)^k y_k$ into Eq. (21·2) and equating

the coefficients of the two lowest powers of $h/2\pi i$ to zero we obtain the following formulas for the first two terms of the expansion:

$$y_0 = \pm [2\mu(E - V)]^{\frac{1}{2}} = \pm p(x,E), \tag{21·3}$$

$$y_1 = -\frac{y_0'}{2y_0} = \frac{d}{dx}[\log p^{-\frac{1}{2}}] = \frac{V'}{4(E - V)}. \tag{21·4}$$

The two different signs of y_0 yield the following linearly independent approximate solutions of Eq. (18·2):

$$f_u(x,E) = p(x,E)^{-\frac{1}{2}} e^{iw(x,E)}, \qquad f_v(x,E) = p(x,E)^{-\frac{1}{2}} e^{-iw(x,E)}. \tag{21·5}$$

$$w(x,E) = \frac{2\pi}{h} \int_{x'}^{x} p(\xi,E)d\xi. \tag{21·6}$$

These functions are usually referred to as B. W. K. or W. B. K. approximations. The exponents are here written as definite integrals to avoid ambiguity regarding the constants of integration. We assume that the potential energy $V(x)$ has the form indicated in Fig. 4 of Sec. 18, and identify the lower limit of integration in (21·6) with the point $x'(E)$ which forms the left-hand boundary of the region of classical motion G for the energy in question (cf. Fig. 4). In the region F to the left of x' the classical local momentum $p(x,E)$ is imaginary, whereas in G it is real. At x' itself, p vanishes.

In G the approximations f_u and f_v, when multiplied by the time factor $e^{-\frac{2\pi i E t}{h}}$, yield progressive waves traveling to the right and left, respectively. If γ is an arbitrary real number, the combination

$$f = e^{i\gamma} f_u + e^{-i\gamma} f_v = 2p^{-\frac{1}{2}} \cos\left(\frac{2\pi}{h} \int_{x'}^{x} p\,dx + \gamma\right) \tag{21·7}$$

represents in the same way a real oscillatory standing wave qualitatively similar in its behavior to the real integral curves $\psi(x)$ of Eq. (18·2).

In F the exponential factors of f_u and f_v are both real and we have only to multiply the approximation functions by $(-1)^{\frac{1}{4}} = e^{i\pi/4}$ in

---

[1] This approximation is the one-dimensional form of that described by Eqs. (11·2), (11·5), and (11·7).

[2] In order to avoid ambiguity due to the multiple-valued character of $p(x,E)$ and $p(x,E)^{\frac{1}{2}}$ a sign convention is needed. We shall have occasion to make use of complex

order to convert them into real functions.   We assume that

$$\lim_{x \to -\infty} iw(x,E) = \frac{2\pi}{h} \int_{-\infty}^{x'} |\sqrt{2\mu(V-E)}| d\xi = \infty.$$

It follows from the sign convention of the footnote that $\lim_{x \to +\infty} iw = \infty$.
Then, clearly, f_u becomes infinite at $x = \mp \infty$, whereas $\lim_{x \to \mp \infty} f_v(x) = 0$.

At the point x' itself, both f_u and f_v become infinite and in the neighborhood of this point they have no value as approximations to $\psi(x)$. This same remark applies equally to the neighborhood of the point x'' which bounds G on the right.

21b. Application to Eigenvalue Problem.—In order to discuss the eigenvalue-eigenfunction problem with the aid of the above approximations it is necessary that we shall be able to determine what particular linear combination of f_u, f_v will fit an exact wave function $\psi(x,E)$ in the regions G and H when the linear combination appropriate to F is known. As the approximations blow up near x' and x'', these neighborhoods form gaps to be bridged.   When appropriate connection formulas are established, it will be possible to use f_u and f_v to evaluate the approximate energies at which the boundary conditions at $x = \pm \infty$ can be fulfilled.

Let us consider first the problem of bridging the gap at $x = x'$. We assume that for large negative values of x the linear combination $\alpha_u f_u + \alpha_v f_v$ describes a real integral curve $\psi(x,E)$ to a satisfactory degree of approximation.   We desire to determine a corresponding linear

---

values of the argument x and hence require a branch of each function which is single-valued over the entire complex plane.   The function $V(x)$ may be assumed to be analytic over the entire plane. $p(x,E)$ will then be analytic over a Riemann surface of two sheets with branch points at x', x'', and at every other simple zero of $V(x) - E$. If the Riemann surface is cut along the axis of reals from x' through x'' to $+\infty$ and `from every other branch point along a radial line leading to infinity, it will separate into two sheets over each one of which $p(x,E)$ is single-valued.   The functions $p^{\frac{1}{2}}$, f_u f_v have the same branch points as p.   Hence similar cuts in the corresponding Riemann surfaces yield single-valued branches of all the functions in which we are interested.

Let us now pick out that branch of p, say p_1, which is positive imaginary along the uncut axis of reals in F, and that branch of $p^{\frac{1}{2}}$, say $(p^{\frac{1}{2}})_1$, which makes $e^{i\pi/4}p^{-\frac{1}{2}}$ a positive real in F.   We identify $w(x,E)$, f_u, and f_v uniquely by postulating that in formulas (21·5) and (21·6) p and $p^{\frac{1}{2}}$ are to be interpreted as p_1 and $(p^{\frac{1}{2}})_1$, the path of integration never being permitted to cross any of the cuts.

When applying these formulas to points on the axis of reals in G and H we shall always use those values appropriate to the upper lip of the cut unless the contrary is explicitly stated.

It follows from these conventions that on the upper lip of the cut in G, p and $p^{\frac{1}{2}}$ are positive real.   On the upper lip in H, p is negative imaginary and $e^{i\pi/4}p^{-\frac{1}{2}}$ is a positive real.

combination $\beta_u f_u + \beta_v f_v$ which shall give a satisfactory description of the same integral curve $\psi(x,E)$ in that portion of G which is remote from x' and x''.

This problem was attacked and partially solved by Jeffreys (and independently by Kramers) using a linear approximation, say

$$E - k(x - x'),$$

for $V(x)$ in the neighborhood of the point x'. Equation (18·2) is thereby reduced to a form which can be integrated exactly, yielding Bessel's functions of order $\frac{1}{3}$ for $\psi(x)$. If the true potential function does not differ from the approximate one in the region where f_u and f_v are bad, the exact solutions for the modified potential can be used to bridge the gap and to give a description of $\psi(x)$ in the neighborhood of x'.

In this way it is possible to derive the Kramers connection formula

$$e^{\frac{i\pi}{4}} f_v = |p|^{-\frac{1}{2}} e^{-\frac{2\pi}{h} \int_x^{x'} |p(\xi)| d\xi} \rightarrow 2p^{-\frac{1}{2}} \cos\left[\frac{2\pi}{h}\int_{x'}^x p(\xi) d\xi - \frac{\pi}{4}\right] \quad (21·8)$$

$$\underbrace{\hspace{4.5cm}}_{\text{Region F}} \qquad \underbrace{\hspace{5cm}}_{\text{Region G}}$$

A proof of the relations (21·8) and (21·9) will be given in Sec. 21e. The sign $\rightarrow$ indicates that if the left-hand member is fitted to ψ in F, the right-hand member will be fitted to ψ in G, but that the reverse is not necessarily true. A similar connection formula

$$2p^{-\frac{1}{2}} \cos\left[\frac{2\pi}{h}\int_{x}^{x''} p(\xi) d\xi - \frac{\pi}{4}\right] \leftarrow |p|^{\frac{1}{2}} e^{-\frac{2\pi}{h}\int_{x''}^{x} |p(\xi)| d\xi} \quad (21·9)$$

$$\underbrace{\hspace{5cm}}_{\text{Region G}} \qquad \underbrace{\hspace{4cm}}_{\text{Region H}}$$

applies at the right-hand boundary. Putting these together we see that if $\psi(x)$ is an eigenfunction of the differential equation (18·2), *i.e.*, if it is a solution which vanishes exponentially at the boundary points, the approximate expressions for $\psi(x)$ in the *good portion of the region G* given by (21·8) and (21·9) should be equal. Accordingly

$$2p^{-\frac{1}{2}} \cos\left[\frac{2\pi}{h}\int_{x'}^x p(\xi) d\xi - \frac{\pi}{4}\right] = 2Cp^{-\frac{1}{2}} \cos\left[\frac{2\pi}{h}\int_{x''}^x p(\xi) d\xi + \frac{\pi}{4}\right].$$

Here C is a real constant which may be either positive or negative. The equation can be valid for a finite interval in G, only if the arguments of the two cosine functions differ by an integral multiple of π. Consequently the condition for an eigenvalue is that

$$\frac{2\pi}{h}\left[\int_{x'}^x p d\xi - \int_{x''}^x p d\xi\right] - \frac{\pi}{2} = n\pi, \qquad n = 0, 1, 2, \cdots$$

or[1]

---

[1] Equation (21·10) is equivalent to the equation

$$\oint\left(y_0 + \frac{h}{2\pi i}y_1\right)d\xi = nh, \qquad n = 0, 1, 2, 3, \cdots$$

where y_0 and y_1 are defined by (21·3) and (21·4).

$$J(E) \equiv \oint p(E,\xi)d\xi \equiv 2\int_{x'}^{x''} p(E,\xi)d\xi = (n + \tfrac{1}{2})h. \qquad n = 0, 1, 2, \cdots$$

$$(21\cdot10)$$

The left-hand member is the familiar Sommerfeld phase integral taken over a complete classical oscillation and Eq. (21·10) is therefore the Bohr theory quantum condition with "half-integral" vibrational quantum numbers. In this approximation the wave theory and the Bohr theory have the same eigenvalues.

In applying the above formula one evaluates the integral $J(E)$ with the aid of the explicit expression for $V(x)$ appropriate to the problem in hand and then solves (21·10) for E in terms of n.

Convenient and simple as the formulas (21·8), (21·9), and (21·10) are, it is necessary to use them with care to avoid mistakes. This means that a careful consideration of their derivation and of the errors involved is important. To this task we now address ourselves, using, however, in place of the Jeffreys-Kramers method a scheme of attack first applied to this problem by Zwaan[1] and more fully developed by the author.[2]

*21c. Zwaan's Method and the Stokes Phenomenon.—Zwaan establishes a connection between the regions to the left and right of x' by introducing complex values of the independent variable and passing around x' in the complex x plane. Any linear combination of f_u and f_v conforms to the differential equation

$$f'' + \left[\frac{4\pi^2 p^2}{h^2} - Q\right]f = 0, \qquad (21\cdot11)$$

where

$$Q = \frac{3}{4}\left(\frac{p'}{p}\right)^2 - \frac{p''}{2p} = \frac{1}{4}\left[\frac{5}{4}\left(\frac{V'}{E-V}\right)^2 + \frac{V''}{E-V}\right]. \qquad (21\cdot12)$$

[1] A. ZWAAN, *loc. cit.*, and *Thesis*, "Intensitäten im Ca-Funkenspektrum," Utrecht, 1929. The discussion by Zwaan (a pupil of Kramers) is an adaptation to the problem under consideration of the previous work of Sir George Stokes. *Cf.* G. Stokes, *Math. and Phys. Papers*, Vol. IV, pp. 77–109, 283–298.

Still another powerful procedure is due to Langer. It is based on the introduction of modified approximate solutions of (18.2) having the form

$$v = \left(\frac{h}{2\pi}\right)^{1/6}\left(\frac{w}{p}\right)^{1/2}J_{\pm 1/3}(w),$$

in which $J_{\pm 1/3}(w)$ denotes a Bessel's function of order $\pm\frac{1}{3}$ in the argument $w(x,E)$ defined by Eq. (21·6). This type of approximation degenerates into a B.W.K.-type function for large values of w, but remains finite when w vanishes. *Cf.* R. E. LANGER, *Trans. Am. Math. Soc.* **33**, 23 (1931), **34**, 447 (1932); *Bull. Am. Math. Soc.* **40**, 545 (1934); *Phys. Rev.* **51**, 669 (1937).

[2] E. C. KEMBLE, *Phys. Rev.* **48**, 549 (1935).

Comparing this with Eq. (18·2) which can be rewritten in the form

$$\psi'' + \frac{4\pi^2 p^2}{h^2}\psi = 0, \tag{21·13}$$

we see that if Q is very small in comparison with $4\pi^2 p^2/h^2$ the equations are practically the same. Thus f is to be regarded as a good approximate solution of (21·13) in the neighborhood of a point x if $\dfrac{h^2}{4\pi^2}\left|\dfrac{Q(x)}{p(x)^2}\right| \ll 1$. If it is possible to find a path Γ in, say, the upper half of the complex plane, joining real points to the left and right of x', crossing none of the cuts mentioned in footnote 2, p. 92, and having the property that

$$\frac{h^2}{4\pi^2}\left|\frac{Q(x)}{p(x)^2}\right| \ll 1$$

along its entire length, we might be led to expect that the approximation f would cling to the corresponding exact function ψ along the entire path. On this hypothesis the coefficients β_u, β_v would be equal to α_u and α_v, respectively. If $E - V$ has a large enough maximum at the potential-energy minimum[1] and if the function $V(x)$ is smooth enough in the neighborhood of $x = x'$, a path Γ satisfying the above inequality can be found. On the other hand, the inference that f must therefore cling to ψ and that β_u and β_v must be equal to α_u and α_v, respectively, is certainly false, for it follows from the assumed reality of $\psi(x)$ on the real axis that $\beta_v = \beta_u{}^*$, whereas $\alpha_u e^{-\frac{i\pi}{4}}$, $\alpha_v e^{-\frac{i\pi}{4}}$ are real. Thus the reality condition is incompatible with the assumption that the same linear combination of f_u and f_v can represent a real integral curve $\psi(x)$ on both sides of x'. The difficulty is of course connected with the fact that $\psi(x)$ is analytic at x', while f_u and f_v have branch points there.

We conclude that if $f(\xi,x)$ denotes that linear combination of $f_u(x)$ and $f_v(x)$ which fits $\psi(x)$ best at the point ξ, the coefficients a_u, a_v in the identity

$$f(\xi,x) \equiv a_u(\xi)f_u(x) + a_v(\xi)f_v(x) \tag{21·14}$$

must change as ξ moves around the curve Γ.[2] The essential feature of the situation to which Zwaan calls attention is the fact that along a

---

[1] In the special case where $V(x) = \frac{1}{2}kx^2$ (harmonic oscillator), the value of $\dfrac{h^2}{4\pi^2}\left|\dfrac{Q}{p}\right|$ at the minimum point $x = 0$ is $h^2\nu^2/8E^2$, where ν is the classical vibration frequency. The value of $\dfrac{h^2}{4\pi^2}\left|\dfrac{Q}{p^2}\right|$ at the mid-point of G for the fifth energy level $[E = (4 + \frac{1}{2})h\nu]$ works out to be $\frac{1}{162}$.

[2] This phenomenon is characteristic of approximations based on asymptotic series and is called the Stokes phenomenon. *Cf.* literature cited in footnote 4, p. 91.

portion of the path Γ the function f_u becomes very small, while f_v becomes very large. Hence f_v and f_u are said to be *dominant* and *subdominant*, respectively, in the region in question.[1] Thus there is no real contradiction between the statement that a *fixed* linear combination of f_u and f_v is a good approximate solution of (21·13) in the neighborhood in question, and the paradoxical requirement that if $a_u f_u + a_v f_v$ is to be an exact solution of (21·13) the coefficient a_u must change markedly in the neighborhood in question.

Zwaan *assumes* that the coefficient of the subdominant term, *i.e.*, a_u, alone can change on Γ. Then we must have $\beta_v = \alpha_v$ and the reality condition fixes β_u as equal to $\alpha_v{}^*$. This result leads at once to the connection formula (21·8) as shown in Sec. 21e below.

*21d. Analysis of the Stokes Phenomenon.—Zwaan's argument is fundamentally sound, but it is not perfectly rigorous and gives little idea of the possible errors involved in the resulting formulas. Hence we shall here undertake a somewhat more detailed analysis of the situation.

The first step in such an analysis must be the laying down of unique definitions of a_u and a_v. We assume that the linear combination

$$a_u f_u + a_v f_v$$

is to be fitted to an exact solution of (21·13), say $\psi(x)$. The fit is to be rigorous except for the isolated singularities of f_u and f_v. Then

$$a_u(x)f_u(x) + a_v(x)f_v(x) = \psi(x). \tag{21·15}$$

This equation is insufficient, however, to fix the values of both coefficients and must be supplemented by another. As a second relation we impose the requirement that

$$a_u(x)f_u'(x) + a_v(x)f_v'(x) = \psi'(x). \tag{21·16}$$

Equations (21·15) and (21·16) are evidently the equations we should have to employ if we were fitting a *fixed* linear combination to ψ at the point x and wished the combination to cling to ψ as closely as possible in the neighborhood of x. The functions a_u and a_v thus defined reduce to α_u and α_v, respectively, at one end of the path Γ and to β_u and β_v, respectively, at the other. As the so-called Wronskian determinant $f_u f_v' - f_u' f_v$ has the constant value $-\dfrac{4\pi i}{h}$, it is always possible to solve Eqs. (21·15) and (21·16) for a_u and a_v. We obtain

$$a_u = \frac{ih}{4\pi}(\psi f_v' - \psi' f_v); \quad a_v = -\frac{ih}{4\pi}(\psi f_u' - \psi' f_u). \tag{21·17}$$

Differential equations for a_u and a_v are easily set up. Thus

$$\frac{da_u}{dx} = \frac{ih}{4\pi}(\psi f_v'' - \psi'' f_v).$$

[1] *Cf.* Appendix D, Part I.

On reduction with the aid of (21·11) and (21·13),

$$\frac{da_u}{dx} = \frac{ih}{4\pi}Qf_v\psi = \frac{ih}{4\pi}\frac{Q}{p}(a_u + e^{-2iw}a_v).$$ (21·18)

Similarly,

$$\frac{da_v}{dx} = -\frac{ih}{4\pi}Qf_u\psi = -\frac{ih}{4\pi}\frac{Q}{p}(a_v + e^{2iw}a_u).$$ (21·19)

Introducing the fixed point ζ, let us make the transformation

$$b_u(x,\zeta) = a_u(x)e^{-iF(x,\zeta)}; \qquad b_v(x,\zeta) = a_v(x)e^{-iF(x,\zeta)}$$

suggested to the author by Dr. Eugene Feenberg. Here $F(x,\zeta)$ is defined

by the equation $F(x,\zeta) = \dfrac{h}{2\pi}\displaystyle\int_\zeta^x \frac{Q}{p}dx$. We obtain the simplified equations

$$\frac{db_u}{dx} = \frac{ih}{4\pi}\frac{Q}{p}e^{-2i(F+w)}b_v;$$

$$\frac{db_v}{dx} = -\frac{ih}{4\pi}\frac{Q}{p}e^{2i(F+w)}b_u.$$

It is convenient at this point to introduce a quantity μ_Γ which we shall call the *index of quality* of a path Γ, and which we define as a line integral along Γ. Thus, let

$$\mu_\Gamma = \frac{h}{2\pi}\int_\Gamma \left|\frac{Q}{p}\right|ds = \frac{h^2}{4\pi^2}\int_\Gamma \left|\frac{Q}{p^2}\right||dw|,$$ (21·20)

where $ds = |dx|$ denotes an element of arc. *We further define a good path Γ as one which satisfies the condition $\mu_\Gamma \ll 1$.* For points x,ζ which can be connected by such a path $|F(x,\zeta)| \ll 1$. Under these circumstances a_u, a_v are sensibly equal to b_u, b_v, respectively. Hence we can write

$$\frac{da_u}{dx} = \frac{ih}{4\pi}\frac{Q}{p}e^{-2iw}a_v, \qquad \frac{da_v}{dx} = -\frac{ih}{4\pi}\frac{Q}{p}e^{2iw}a_u$$ (21·21)

in discussing the behavior of a_u and a_v along a path which is good and can be connected with ζ by an extension which is also good.

Let us apply the term *Stokes region* to a portion of the complex x plane on which $\dfrac{h}{2\pi}\left|\dfrac{Q}{p}\right|$ and $\dfrac{h^2}{4\pi^2}\left|\dfrac{Q}{p^2}\right|$ are very small and throughout which $|e^{iw}|$ is either very large or very small. If $|e^{iw}|$ is very large, $|f_u| \ll |f_v|$ and f_u is dominant. If $|e^{iw}|$ is very small, f_v is dominant. We call these portions of the complex plane Stokes regions because, as we shall see, they are the domains in which the Stokes phenomenon of the shifting of the coefficients takes place. (Changes in these coefficients also occur near the

classical turning points x' and x'', but such changes are of no interest here because no good paths can be drawn through these neighborhoods and the differential equations (18·2) and (21·21) are in violent disagreement there.) Figure 6 shows the Stokes regions for the harmonic oscillator problem. A and B are regions in which f_u is dominant, while C and D are regions in which f_v is dominant. A modification of the potential-energy function implies the introduction of additional real or complex zeros of $E - V(x)$, but, if these zeros are not too close to x' and x'', the Stokes regions near these points will remain substantially as indicated in the figure for the harmonic oscillator.
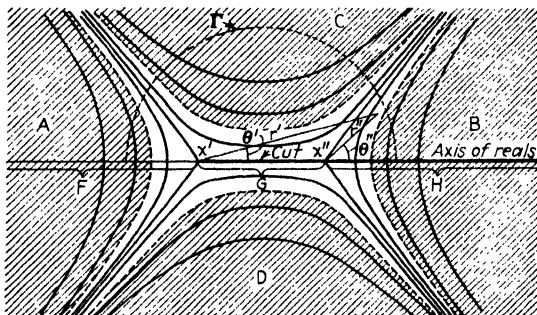
Fig. 6.—Level lines of the function $|e^{iw}|$ and Stokes regions for the harmonic oscillator problem.

Our problem is to determine the coefficients β_u, β_v in terms of α_u, α_v by means of an approximate integration of Eqs. (21·21) along a suitably chosen good path Γ. To this end it is convenient to consider, separately, several different types of good path.

First of all, let us deal with a path which lies wholly in one of the "neutral zones" between the different Stokes regions. On such a path, $|e^{iw}|$ is of the order of unity and it follows from Eqs. (21·21) that a_u and a_v are sensibly constant.

Consider next a Stokes region M in which f_u is dominant. According to the above equations da_u/dx is very small in M, and we infer that a_u can be treated as sensibly constant along any good path Γ confined wholly to this region. On the other hand, a_v is evidently far from constant in M except in the special case where $a_u(x)$ is zero or nearly zero at the initial point of the path. In that case $a_u(x)$ must be nearly equal to zero along the entire length of the path and the second of Eqs. (21·21) indicates that a_v may be sensibly constant.

In the case of a path located in a Stokes region in which f_v is dominant the behavior of the coefficients a_u and a_v is reversed. Thus our conclusion is that *each coefficient changes only in Stokes regions in which the corresponding function f_u or f_v, as the case may be, is subdominant. Even the coefficient*

of the subdominant term will be sensibly constant along a good path, however, if the coefficient of the dominant term is zero, or sufficiently small.

In order to put these statements on a rigorous basis and determine their limits of error, the writer has made a more exact analysis of the situation.[1] Consider first the variation in a_u along a path Γ leading from the initial point x_0 to the final point x_1 and *so drawn that $|e^{iw}|$ increases steadily as we move from the initial point toward the final point.* In this case the inequality

$$|a_u(x_1) - a_u(x_0)| < \tfrac{1}{2}\mu_\Gamma[|a_u^0| + |a_v^0||e^{-2iw_0}|] \tag{21·22}$$

can be established. If the path starts in a neutral zone where $|e^{-2iw_0}|$ is nearly equal to unity, we have

$$|a_u(x_1) - a_u(x_0)| < \tfrac{1}{2}\mu_\Gamma[|a_u^0| + |a_v^0|]. \tag{21·23}$$

The variation in a_v along a similar good path is limited by the inequality

$$|a_v(x_1) - a_v(x_0)| < \tfrac{1}{2}\mu_\Gamma[|a_v^0| + |a_u^0||e^{2iw_0}|], \tag{21·24}$$

where, however, x_0 is now taken to be at the uphill end of the path where f_u is dominant. The inequalities (21·22) and (21·24) readily confirm the italicized statements of the preceding paragraph, provided that we understand that a coefficient is to be regarded as sensibly constant along a path when its change is small compared with the largest of the four quantities $|a_u(x_0)|$, $|a_v(x_0)|$, $|a_u(x_1)|$, $|a_v(x_1)|$.

*21e. Derivation of the Connection Formulas.—Having examined the general problem of the variation in the a's, we return to the problem of the connection formulas. We desire to determine the coefficients β_u, β_v for the good portion of the axis of reals in the region G in terms of the coefficients α_u, α_v in F. To make the problem quite definite we choose a pair of definite points, say P_1 in F and P_2 in G, at which the two pairs of coefficients are to be evaluated. In order to set up the desired formulas we must assume the existence of a path Γ joining P_1 with P_2 through the upper half plane and having the property that $\mu_\Gamma \ll 1$. This path must cross none of the cuts on which f_u and f_v are discontinuous and hence must enclose no zeros of $E - V(x)$ between itself and the portion of the axis of reals between P_1 and P_2. It will necessarily pass through the Stokes region C in which f_v is dominant, as well as through part of the Stokes region A in which it originates.

Consider first the case of a ψ function which vanishes at $x = -\infty$. Then $\alpha_u = a_u(P_2) = 0$ and both coefficients are constant in A. a_v remains constant along that portion of Γ which traverses C and we conclude that $\beta_v = \alpha_v$. On the other hand, the coefficient a_u is not constant in C and we should have to perform a difficult task in integration in order to determine β_u from the differential equation for this coefficient. Fortunately the integration is unnecessary as we can fall back on the

[1] *Cf.* footnote 2, p. 95.

reality condition used on p. 96. Giving α_v the value $e^{i\pi/4}$ in order to make ψ real in F, it follows that

$$\beta_u = \beta_v{}^* = \alpha_v{}^* = e^{-\frac{i\pi}{4}}.$$

Thus the connection formula (21·8) is established.

As stated on p. 94, we draw the arrow in (21·8) from left to right to indicate that the approximate validity of the left-hand member implies that of the right, whereas the converse statement is not true. In order to make clear the reason for this "one-way street sign" and prepare the way for the derivation of the second connection formula, we observe that in view of the homogeneous linear character of Eqs. (21·18) and (21·19), the coefficients β_u and β_v must be homogeneous linear functions of α_u and α_v. Thus a complete exact solution of the whole connection problem would involve a complete exact determination of the matrix $\|g\|$ of the equations

$$\left. \begin{aligned} \beta_u &= g_{uu}\alpha_u + g_{uv}\alpha_v; \\ \beta_v &= g_{vu}\alpha_u + g_{vv}\alpha_v. \end{aligned} \right\} \tag{21·25}$$

The approximate solution of Eqs. (21·18) and (21·19) for the special case that $\alpha_u = 0$ shows that

$$g_{vv} = 1 + \delta; \qquad g_{uv} = -i(1 + \delta^*),$$

where δ is a small quantity of the order of magnitude of μ_Γ which we have neglected in Eq. (21·8).

Additional information regarding the matrix $\|g\|$ is obtainable from the fact that the Wronskian of any two exact solutions of the Schrödinger equation, say ψ and $\bar\psi$, is constant along the axis of reals. Thus, if

$$\psi = \alpha_u f_u + \alpha_v f_v, \qquad \bar\psi = \bar\alpha_u f_u + \bar\alpha_v f_v$$

in F, it follows that, in F,

$$W[\psi,\bar\psi] = \psi\frac{\partial\bar\psi}{\partial x} - \bar\psi\frac{\partial\psi}{\partial x} = \frac{4\pi i}{h}(\alpha_v\bar\alpha_u - \alpha_u\bar\alpha_v) = \text{constant}.$$

The Wronskian of the same pair of solutions in G takes the form

$$W[\psi,\bar\psi] = \frac{4\pi i}{h}(\alpha_v\bar\alpha_u - \alpha_u\bar\alpha_v)(g_{vv}g_{uu} - g_{uv}g_{vu}).$$

Equating these two expressions for the Wronskian, we see that the determinant of $\|g\|$ is unity. Hence the inverse equations to (21·25) are

$$\left. \begin{aligned} \alpha_u &= g_{vv}\beta_u - g_{uv}\beta_v, \\ \alpha_v &= -g_{vu}\beta_u + g_{uu}\beta_v. \end{aligned} \right\} \tag{21·26}$$

It follows from the reality condition that if $\beta_u = \beta_v{}^*$, the quantities $\alpha_u e^{-i\pi/4}$ and $\alpha_v e^{-i\pi/4}$ are real. Hence

$$g_{vu} = -ig_{uu}{}^*, \qquad g_{uv} = -ig_{vv}{}^*. \tag{21·27}$$

Combining the last relation with the requirement that the determinant of $\|g\|$ be equal to unity, we obtain the following relation between g_{uu} and δ:

$$g_{uu}(1 + \delta) + g_{uu}{}^*(1 + \delta^*) = 1. \tag{21·28}$$

If small quantities of the order of μ_Γ be neglected in comparison with unity, one can conclude that the real part of g_{uu} is $\tfrac{1}{2}$. The imaginary part cannot be determined even approximately without an explicit evaluation of a difficult integral, but it is possible with the aid of (21·24) to set up the useful upper bound

$$|g_{uu}| = |g_{vu}| < \mu_\Gamma |e^{2iw(P_1)}| = \mu_\Gamma \exp\left[\frac{4\pi}{h}\int_{P_1}^{z'} |p|d\zeta\right]. \tag{21·29}$$

Let us now consider the validity of the inverse of the relation (21·8) We know from the derivation that if the left-hand member fits the function $\psi(z)$ exactly at P_1 the right-hand member fits the same function approximately in the neighborhood of P_2. The inverse relation would be the statement that if the right-hand member fits $\psi(z)$ exactly at P_2, the left-hand member fits $\psi(z)$ approximately in the neighborhood of P_1. With the aid of our partial determination of the matrix $\|g\|$ we can test the validity of this statement. We accordingly assume, as in the right-hand member of (21·8), that $\beta_u = i\beta_v = e^{i\pi/4}$. It follows that the corresponding values of α_v and α_u are not $e^{i\pi/4}$ and 0, but

$$e^{i\pi/4}[1 - \delta^* + g_{uu}(\delta^* - \delta)]$$

and $(\delta - \delta^*)e^{-\frac{i\pi}{4}}$. Although the correction to α_u is small, its product by the dominant approximation function f_u is not necessarily small. Hence the left-hand member of (21·8) may be entirely incorrect when the right-hand member fits $\psi(z)$ exactly at P_2.

A second connection formula is a direct consequence of (21·26) and our information regarding $\|g\|$. Consider an exact solution of (18·2) which has the form

$$\psi(x) = 2p^{-1/2}\cos(w + \gamma) = e^{i\gamma}f_u + e^{-i\gamma}f_v$$

in the neighborhood of P_2. By (21·26) the corresponding values of α_u and α_v are, if we neglect small quantities of the order of μ_Γ,

$$\alpha_u = 2e^{\frac{i\pi}{4}}\cos\left(\gamma - \frac{\pi}{4}\right), \qquad \alpha_v = e^{\frac{i\pi}{4}}\left[2X\cos\left(\gamma - \frac{\pi}{4}\right) - \sin\left(\gamma - \frac{\pi}{4}\right)\right].$$

Here the symbol X stands for the unknown, but bounded, imaginary part of g_{uu}. It is not difficult to show from these formulas and the

inequality (21·29) that if $\mu_\Gamma \ll 1$ and $\left|\tan\left(\gamma - \dfrac{\pi}{4}\right)\right| \ll 2|e^{2iw(P_1)}|$, the product $\alpha_v f_v$ is small compared with $\alpha_u f_u$ in the neighborhood of P_1. Thus we obtain the connection formula

$$\underbrace{\cos\left(\gamma - \frac{\pi}{4}\right)|p|^{-\frac{1}{2}}e^{\frac{2\pi}{h}\int_x^{x'}|p|d\xi}}_{\text{Region F}} \leftarrow \underbrace{p^{-\frac{1}{2}}\cos\left[\frac{2\pi}{h}\int_{x'}^x pd\xi + \gamma\right].}_{\text{Region G}} \quad (21\cdot30)$$

This is usually specialized by setting γ equal to $\pi/4$ in which case the condition laid upon $\tan\left(\gamma - \dfrac{\pi}{4}\right)$ is automatically fulfilled for all positions of P_1. The possibility of the above generalization to values of γ different from $\pi/4$ has been indicated by Langer.[1]

In addition to the restriction on γ, the sole condition for the validity of the connection formulas (21·8) *and* (21·30) *is that there shall exist a path* Γ *in the upper half plane connecting* P_1 *with* P_2, *enclosing no complex zeros of* $E - V(x)$, *and having the property that* $\mu_\Gamma \ll 1$.

The relations (21·8) and (21·30) apply to the left-hand boundary x' of the classical region G, where $V(x)$ has a negative slope. The corresponding relations for the right-hand boundary x'' are (21·9) and

$$\underbrace{p^{-\frac{1}{2}}\cos\left[\frac{2\pi}{h}\int_x^{x''}pd\xi + \gamma\right]}_{\text{Region G}} \rightarrow \underbrace{|p|^{-\frac{1}{2}}\cos\left(\gamma - \frac{\pi}{4}\right)e^{\frac{2\pi}{h}\int_{x''}^x|p|d\xi}}_{\text{Region H}}. \quad (21\cdot31)$$

The last formula applies to a pair of points P_3 in G and P_4 in H, provided that they can be connected by the usual good path Γ and that

$$\left|\tan\left(\gamma - \frac{\pi}{4}\right)\right| \ll 2|e^{2iw(P_4)}|.$$

*21f. Derivation of the Sommerfeld Phase-integral Quantum Condition.—The Sommerfeld phase-integral quantum condition can be derived with the aid of the connection formulas (21·8) and (21·9). We assume that Eq. (21·13) is to be solved in the infinite interval

$$-\infty < x < +\infty$$

with the boundary conditions $\lim_{x \to \pm \infty} \psi(x) = 0$, the potential function being everywhere differentiable and greater than the energy E except in the single region of classical motion G. It follows that a_u must vanish at $x = \pm \infty$. By (21·18) $a_u = \dfrac{ih}{4\pi}\displaystyle\int_{-\infty}^{P_1} Qf_v\psi dx$ with the integration carried out along the axis of reals.

[1] *Loc. cit.*, footnote 1, p. 95.

Obviously, then, α_u will not be exactly zero and we must make sure that it is small enough to make (21·8) applicable. Replacing x by w as the variable of integration we obtain the upper bound

$$|\alpha_u| \leqslant \frac{1}{2} \int_{w(-\infty)}^{w(P_1)} \left| \frac{Qh^2}{4\pi^2 p^{3_2}} \right| e^{-|w|} |\psi| |dw|.$$

Let N denote the maximum value of $\left| \frac{Qh^2}{4\pi^2 p^{3_2}} \right|$ in the range of integration. Since $|\psi(x)|$ decreases monotonically as we pass from P_1 out to $-\infty$ along the axis of reals (*cf.* Appendix C), we may replace it by $|\psi(P_1)|$. Thus

$$|\alpha_u| < \frac{N}{2} |\psi(P_1)| \int_{w(P_1)}^{w(-\infty)} e^{-|w|} d|w|$$

$$< \frac{N}{2} |\psi(P_1)| e^{-|w(P_1)|}$$

$$< \frac{N}{2|p(P_1)^{1/2}|} [|\alpha_u| + |\alpha_v| e^{-2|w(P_1)|}]. \tag{21·32}$$

If f_v is a good approximate solution to (21·13) throughout the range of integration, $N/|p(P_1)^{1/2}|$ will be much less than unity and

$$|\alpha_u| \ll |\alpha_v| e^{-2 w(P_1)}.$$

Let $\delta\beta_u$ and $\delta\beta_v$ denote the errors in β_u and β_v due to setting $\alpha_u = 0$, as in the left-hand member of (21·8). Then

$$|\delta\beta_u| = |g_{uu}| |\alpha_u|; \qquad |\delta\beta_v| = |g_{vu}| |\alpha_u|.$$

The inequality (21·29) gives an upper bound on $|g_{uu}|$ and $|g_{vu}|$, while the "unperturbed" values of $|\beta_u|$ and $|\beta_v|$ are both equal to $|\alpha_v|$ if we neglect small quantities of the order of μ_Γ. Hence

$$\frac{|\delta\beta_u|}{|\beta_u|} < \frac{N\mu_\Gamma}{2|p(P_1)^{1/2}|}; \qquad \frac{|\delta\beta_v|}{|\beta_v|} < \frac{N\mu_\Gamma}{2|p(P_1)^{1/2}|}.$$

These inequalities show that the error in using (21·8) for the determination of the form of an eigenfunction in the classical domain G is negligible provided that μ_Γ and $N/2|p(P_1)^{1/2}|$ are small compared with unity. Of course (21·9) stands on the same footing in this respect as (21·8).

Let us now assume that the energy level under consideration is high enough so that there is a portion of the region G in which $\frac{h^2}{4\pi^2} \left| \frac{Q}{p^2} \right| \ll 1$. If there are no complex roots of $E - V(x)$ near the interval G of the axis of reals, we can assume the existence of good paths Γ and Γ', connecting G with a pair of points P_1 in F, and P_4 in H, respectively. Equations (21·8) and (21·9) give two different approximate expressions for $\psi(x)$ in the good portion of G. Equating these expressions, one obtains

$$2p^{-1/2} \cos\left[\frac{2\pi}{h} \int_{x'}^{x} p(\xi) d\xi - \frac{\pi}{4} \right] = 2Cp^{-1/2} \cos\left[\frac{2\pi}{h} \int_{x''}^{x} p(\xi) d\xi + \frac{\pi}{4} \right]$$

as given in Sec. 21*b*. This equation leads at once to the Sommerfeld formula (21·10).

The energy levels given by (21·10) for the ideal linear oscillator are identical with those derived by rigorous methods in Sec. 20. This is somewhat surprising inasmuch as the conditions for the validity of the connection formulas are not fulfilled at the lower energy levels. In the case of the lowest level, for example, the value of $\dfrac{h^2}{4\pi^2}\left|\dfrac{Q}{p^2}\right|$ at the mid-point

of G is $\frac{1}{2}$ and $\dfrac{h}{2\pi}\displaystyle\int\left|\dfrac{Q}{p}\right|ds$ has a minimum value of about $\frac{1}{3}$ for any path leading from the point in question to the Stokes regions B, A, and C (Fig. 6). Consequently we are led to inquire whether there is not some general principle which will permit us to establish Eq. (21·10) independently of the connection formulas (21·8) and (21·9). The answer is that such a principle does exist. We state only the general result in this section, referring the reader to Part II, Appendix D, for the proof.[1]

Let the basic problem under consideration be the same as the one discussed above, *viz.*, that of finding the discrete energy levels for a one-dimensional anharmonic oscillator with the fundamental interval $-\infty < x < +\infty$, an analytic potential function, and a single potential valley. We assume that for the energy or energies under discussion $\dfrac{h^2}{4\pi^2}\left|\dfrac{Q}{p^2}\right|$ is small throughout the intervals F and H except in the immediate neighborhood of the end points x' and x''. Instead of postulating good paths Γ' and Γ'' connecting G with F and H, however, we postulate a single path Γ connecting a pair of points P_1 in F and P_4 in H without enclosing any complex zeros of $E - V(x)$, and having the property that $\mu_\Gamma \ll 1$. We further assume that $N/|p(P_1)^{1/2}|$ is small compared with unity—a parallel condition for the point P_4 would be equally good. *Under these conditions Eq. (21·10) can be shown to hold even for the low eigenvalues for which no good paths connecting G with F and H can be set up.*

The only doubtful point likely to arise in the application of this result to specific problems is the question whether the good path Γ can actually be set up. In order to test this point qualitatively one can replace $E - V(x)$ by an approximating polynomial of degree k,

$$T(x) = \text{const.} \times \prod_{j=1}^{k}(x - x_j).$$

If this approximation is good throughout the interval between suitably chosen points P_1 and P_4 we can compute $\left|\dfrac{Q}{p}\right|$ and $\left|\dfrac{Q}{p^2}\right|$ along a path connecting P_1 with P_4 from $T(x)$. The condition that a path Γ shall be

---

[1] The work of Part II, Appendix D, leads to essentially the same result as that obtained by Birkhoff (*loc. cit.*, footnote 4, p. 91) by entirely different methods.

"good" is then readily seen to be that it shall keep away from all the complex roots of $T(x)$.

The writer has carried through a test computation of the accuracy of (21·10) using the Morse potential function for the normal state of H_2 molecule. This function has the form

$$V = -2De^{-ax} + De^{-2ax},$$

where x is the displacement from the equilibrium position. Let $\zeta = e^{-ax}$ Then

$$E - V = E + 2D\zeta - D\zeta^2.$$

The zeros of this function are

$$\zeta = 1 \pm \left(1 + \frac{E}{D}\right)^{\frac{1}{2}},$$

or

$$x = -\frac{1}{a}\left[\log_e\left(1 \pm \left(1 + \frac{E}{D}\right)^{\frac{1}{2}}\right) + i2\pi m\right], \qquad m = 0, \pm 1, \pm 2, \cdots$$

Thus we obtain two sequences of zeros in the complex x plane equally spaced along lines parallel to the imaginary axis and passing through x' and x''. The computation is simplified by choosing E equal to $-D$ so as to bring the points x' and x'' together. This should not materially affect the value of μ_Γ for a path that cuts midway between the nodes x', x'' and the next pair above them. Taking Γ to be a circle of radius π/a with its center at the origin, and inserting the values of a, D, and the mass coefficient appropriate to the normal state of H_2, we obtain $\mu_\Gamma = 0.015$. The quantity $N/|p(P_1)^{\frac{1}{2}}|$ of (21·32) is so small that errors due to the finite values of a_u at P_1 and P_4 are negligible in comparison with those due to μ_Γ. Let δE denote the maximum error involved in using (21·10) and let ΔE denote the energy-level spacing for adjacent values of the quantum number n. Then $\left|\dfrac{\delta E}{\Delta E}\right|$ can be worked out and proves to be sensibly equal to $\mu_\Gamma/2\pi$. Thus we can trust the Sommerfeld formula in this case to about three-tenths of a per cent of the spacing of adjacent energy levels.

Of course the actual error involved in (21·10) should be appreciably less than the computed upper bound. Moreover, since $\left|\dfrac{Q}{p}\right|$ varies inversely as the square root of the effective mass, the above figure would be materially improved in the case of a heavier molecule. In the case of a harmonic oscillator (quadratic potential function) there are no roots of $E - V(x)$ except those at x' and x''. Hence we can choose Γ as a semicircle of indefinitely large radius and thus reduce the maximum error of the Sommerfeld formula below any assignable limit.

*21g. Higher Approximations.—A better approximation to the energy eigenvalue can be obtained in most cases by using higher order terms in the series expansion of the function y of Eq. (21·2).[1] If we take an n-term approximation

$$y^{(n)}(x) = \sum_{k=0}^{n} \left(\frac{h}{2\pi i}\right)^k y_k(x),\tag{21·33}$$

where n is not too large, the differential equation for the function $e^{\frac{2\pi i}{h}\int_{x'}^{x} y^{(n)} dx}$ will usually be a better approximation to (18·2) along Γ than the approximations we have been using. It follows from the argument of Part II, Appendix D, that the above function must have the same values on the upper and lower lips of the cut in H. Thus we can replace Eq. (21·10) by the more accurate energy level condition

$$\oint y^{(n)} dx = mh. \qquad m = 0, 1, 2, \cdots\tag{21·34}$$

As previously stated, however, the series obtained by letting n become infinite is only semiconvergent. Hence we cannot improve our result indefinitely by making n very large.

*21h. Modification of Method for Radial Motion in Two-particle Problem.[2]— In applying the B. W. K. method to the radial equation of the two-particle problem [Eq. (28·19), p. 150], one meets with an apparent difficulty in that the fundamental region in which the equation to be solved ranges from 0 to ∞ instead of from $-\infty$ to $+\infty$. Moreover, the B. W. K. approximations do not have the right character to fit the exact solution of the differential equation at the left-hand boundary point $r = 0$.[3] Thus there is no good path Γ leading from one boundary to the other because the left boundary is itself "bad." As pointed out by Krame s,[4] however, it is possible to fit the boundary conditions at both ends of the fundamental region if we modify the approximation formulas by the addition of the term $h^2/32\pi^2\mu r^2$ to the potential energy. The added term is negligible except in the immediate neighborhood of the origin, and it is ordinarily possible to find a path Γ leading from $r = 0$ to $r = \infty$ along which the *modified* B. W. K. functions are good approxi-

[1] *Cf.* J. L. Dunham, *Phys. Rev.* **41**, 713–720, 721–731 (1932).

[2] This section amplifies the brief treatment of the same subject previously given by the author in the reference of footnote 2, p. 95. An article by R. E. Langer in the current number of the *Physical Review* (Apr. 15, 1937) contains an independent discussion of the same problem which overlooks the earlier work of the present writer.

[3] Consider, for example, Eq. (28·19) when $l = 0$ and $V(r)$ has a pole of the first order at the origin. In this case both of the B. W. K. approximations f_u, f_r vanish at the origin, whereas *no* exact solution of the differential equation can have a zero at that point.

[4] *Loc. cit.*, footnote 2, p. 91.

mate solutions of (18·2). Under these circumstances Eq. (21·10) gives the energy levels if the momentum $p(r,E)$ is computed from the modified potential function [*cf.* Sec. 28*g*, Eq. (28·23)].

Let us assume that $V(r)$ has the form

$$V(r) = -\frac{A}{r} + \frac{B}{r^2} + \varphi(r), \qquad 0 < r < \infty$$

where $\varphi(r)$ is analytic along the entire positive real axis and at the origin. Let the modified potential function be

$$\Phi(r) = V(r) + \frac{1}{4\kappa r^2}. \qquad \kappa = \frac{8\pi^2\mu}{h^2}$$

The corresponding modified momentum is

$$p^{(m)} = [2\mu(E - \Phi)]^{\frac{1}{2}}.$$

Let the modified B. W. K. approximations $f_u{}^{(m)}$ and $f_v{}^{(m)}$ be computed from $p^{(m)}$ just as f_u and f_v are computed from p. Then the modified approximations are readily proved to be solutions of the equation

$$\frac{d^2 f^{(m)}}{dr^2} + \left(\frac{4\pi^2 p^2}{h^2} - Q^{(m)}\right) f^{(m)} = 0,$$

in which $Q^{(m)}$ is defined by

$$Q^{(m)} = \frac{1}{4}\left[\frac{1}{r^2} + \frac{5}{4}\left(\frac{\Phi'}{E-\Phi}\right)^2 + \frac{\Phi''}{E-\Phi}\right].$$

Comparing the above differential equation with (21·11), we see that $Q^{(m)}$ plays the part formerly played by Q in determining the quality of the modified approximations and in determining the variation in the coefficients when a linear combination of $f_u{}^{(m)}$ and $f_v{}^{(m)}$ is fitted to an exact wave function $\psi(r)$. Whereas Q becomes infinite as $1/r^2$ at the origin, $Q^{(m)}$ becomes infinite there only as $1/r$. $Q^{(m)}/p^{(m)}$ remains finite at the origin and $Q^{(m)}/[p^{(m)}]^2$ vanishes there.

Thus one is led to hope for a good path, leading from the origin to a point on the axis of reals beyond the region G, along which the coefficient a_v is sensibly constant. Investigation shows that, if $\varphi(r)$ is well behaved, there is no difficulty in finding a path of this kind. The validity of Eq. (21·10) when used with the modified momentum formula is a direct consequence of the existence of the good path.

21i. Asymptotic Agreement of Wave Theory and Classical Theory Regarding Position of Particle.—It is interesting to note that in the case of a state of high quantum number the Bohr theory gives us not only the correct energy levels but also the correct value of the probability of finding the particle on any segment Δx of the x axis, provided that Δx is large compared with the de Broglie wave length $\lambda = h/p(x,E)$ or is made equal to an integral number of half wave lengths.[1] In computing this probability on the basis of the semiclassical Bohr theory, one must assume that the phase of the motion is unknown. The probability in

---

[1] *Cf.* J. H. VAN VLECK, *J. Franklin Inst.* **207**, 475 (1929); *Proc. Nat. Acad. Sci.* **14**, 178 (1928).

question ΔF is then equal to the ratio of the time occupied in crossing the segment Δx to the half period of vibration. Denoting the period by T and the classical velocity by v, we have

$$\Delta F = \frac{2}{T} \int_x^{x+\Delta x} \frac{d\xi}{v(\xi,E)} = \frac{2\mu}{T} \int_x^{x+\Delta x} \frac{d\xi}{p(\xi,E)}. \tag{21.35}$$

Computing $\psi\psi^*$ from the B. W. K. approximation

$$\psi(x,E) = A p^{-\frac{1}{2}} \cos [w(x,E) + \gamma],$$

we obtain

$$dF = \frac{A^2}{p} \cos^2 (w + \gamma)$$

for the probability of an infinitesimal interval dx in the region of real classical momentum. To eliminate the fluctuations due to the nodes of ψ, we replace the phase factor $\cos^2 (w + \gamma)$ by its average value in computing the probability of a macroscopic segment Δx. Thus

$$\Delta F = \frac{A^2}{2} \int_x^{x+\Delta x} \frac{d\xi}{p(\xi,E)}. \tag{21.36}$$

As the probability of finding a particle in the region of imaginary classical momentum becomes negligible for large energies, ψ can be normalized by integrating over the region G alone. In this way one obtains

$$1 = \int_{x'}^{x''} \psi\psi^* dx = \frac{A^2 T}{4\mu}. \tag{21.37}$$

Combining (21.37) with (21.36), one obtains the Bohr equation (21.35).

21j. The Transmission of Progressive Matter Waves through a Potential Hill.—A very important aspect of quantum mechanics is that it permits particles of energy E to penetrate a potential hill where $E - V$ is locally negative. We are already familiar with this penetration as exhibited by the fact that in the one-dimensional oscillator problem of Secs. 18 and 19 the wave function Ψ does not reduce to zero at the classical turning points x' and x''. Another consequence of the continuity of the Schrödinger equation through regions of negative kinetic energy is that particles can actually pass through barriers that would be classically impenetrable and come out on the other side. This phenomenon is sometimes called the "tunnel effect" and forms the basis of recent theories of radioactive decay (cf. Sec. 31), of the extraction of electrons from cold metals by intense electric fields,[1] of the predissociation of molecules,[2] etc. It could be treated by means of wave packets which on

---

[1] Cf. R. H. FOWLER and L. NORDHEIM, *Proc. Roy. Soc.* **119**, 173 (1928).

[2] Cf. O. K. RICE, *Phys. Rev.* **34**, 1451 (1929), **35**, 1538 (1930).

striking the potential barrier are split into two parts, one reflected and the other transmitted through the barrier. A simpler procedure, however, is to study the reflection of a steady stream of particles incident on a potential barrier from one side, using the method of nonquadratically integrable wave functions mentioned in Sec. 8. This is the method almost universally employed for the study of reflection and scattering phenomena.

The B. W. K. approximation functions and the Kramers connection formulas have been much used for the approximate evaluation of the transmission and reflection coefficients for the tunnel effect. The
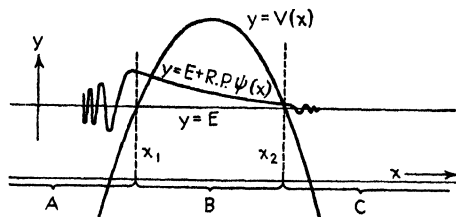
FIG. 7.—The transmission of matter waves through a potential hill. The curve marked $y = E + R.P. \, \psi(x)$ indicates the real part of the wave function for waves of energy E incident on the hill from the left.

approximation functions are particularly adapted to the case of a parabolic potential such as that indicated in Fig. 7. We accordingly assume the potential function $V_0 - \frac{1}{2}kx^2$ and choose the zero level of energy as V_0. The expression for the square of the classical local momentum becomes

$$p^2 = 2\mu(E - V) = \mu k\left(x^2 + \frac{2E}{k}\right).$$

E is assumed to be negative. It is convenient to modify the sign convention of footnote 2, p. 92, and make p negative real and $ip^{-\frac{1}{2}}$ positive real to the left of 'the classical turning point x_1 of Fig. 7. Both p and $p^{-\frac{1}{2}}$ are then positive real to the right of x_2 (cf. Appendix D, Part III). With these conventions f_u represents an outgoing steady current of particles on both sides of the hill and f_v a similar incoming current.

To prove this last statement we introduce the mass current density $\vec{I}$ defined for particles moving in three dimensions by Eq. (8·5), Chap. I. Setting the vector potential of Eq. (8·5) equal to zero, we obtain the expression

$$I = \frac{h}{4\pi i}\left(\psi^*\frac{d\psi}{dx} - \psi\frac{d\psi^*}{dx}\right)$$

for the component of $\vec{I}$ in the positive direction of the x axis in our one-

dimensional problem. The values of I for f_u and f_v on the axis of reals are

$$I[f_u] = \frac{(p + p^*)}{2|p|} e^{i(w-w^*)}; \qquad I[f_v] = -\frac{(p + p^*)}{2|p|} e^{-i(w-w^*)}.$$

In A the functions f_u and f_v give real currents of unit magnitude directed along the negative and positive axes, respectively. Denoting the real integral $\frac{2\pi}{h} \int_{x_1}^{x_2} |p| d\xi$ by K, we find that the current in the region C to the right of x_2 is e^{-2K} for f_u and $-e^{2K}$ for f_v. This verifies the statement at the end of the preceding paragraph.

An examination of the Stokes regions given in Part III, Appendix D, shows that if there is a good path Γ passing around the turning points x_1 and x_2 from A to C without enclosing other roots of $E - V(x)$, we can establish the existence of approximate connection formulas of the character

$$f_u + cf_v \leftarrow f_u; \tag{21.38}$$

$$f_u \rightarrow \hat{f}_u + d\hat{f}_v. \tag{21.39}$$

$$\underbrace{\hspace{3cm}}_{A} \underbrace{\hspace{3cm}}_{C}$$

Here $\hat{f}_u$, $\hat{f}_v$ denote the values of f_u and f_v appropriate to the lower edge of the cut along the portion C of the axis of reals. The upper connection formula is applicable to the reflection of particles incident on the hill from the left, while the lower one applies to a stream of particles incident on the hill from the right.

In the region A to the left of x_1 we have $f_v = -f_u^*$. It follows that if α and β are any constants,

$$I[\alpha f_u + \beta f_v] = I[\alpha f_u] + I[\beta f_v] = |\beta|^2 - |\alpha|^2.$$

Thus the approximate wave function $f_u + cf_v$ represents the superposition of an incident stream of magnitude $|c|^2$ and a reflected stream of unit magnitude. The corresponding transmitted stream in C has the magnitude e^{-2K}. The net current on the two sides of the hill must be the same, as one can prove from the constancy of the Wronskian of any exact wave function ψ and its conjugate ψ^*. Hence

$$|c|^2 - 1 = e^{-2K}. \tag{21.40}$$

The corresponding transmission coefficient for an energy which is less than V_{max} is

$$\theta = \frac{|c|^2 - 1}{|c|^2} = \frac{1}{1 + e^{2K}}. \tag{21.41}$$

A complete determination of the complex constant c would determine the phase of the incident wave with respect to the transmitted and

reflected waves. Unfortunately the methods here employed are not powerful enough for this purpose.[1]

The same transmission coefficient is derivable from (21·39) for particles incident from the right. As pointed out to the writer by Dr. Eugene Feenberg, the equality of the two transmission coefficients for particles incident on either side of an asymmetric potential barrier is a direct consequence of the constancy of the Wronskian of the corresponding pair of wave functions.

Formula (21·41) is free from all restrictions depending on the height of the hill provided that $E - V_{max} < 0$. It holds not only for the ideal case of a parabolic hill but for any analytic potential hill which yields no zeros of $E - V(x)$ near x_1 and x_2 and which therefore permits a good path Γ joining the regions A and C.

If the maximum potential energy is less than the energy of the incident particles, a parabolic potential function yields two imaginary roots of $E - V(x)$, viz.,

$$\left.\begin{array}{c} x_2 \\ x_1 \end{array}\right\} = \pm i \left[\frac{2(E - V_{max})}{-V''} \right]^{1/2}.$$

Denoting the integral $\dfrac{2\pi}{h} \displaystyle\int_{x_1}^{x_2} |p||d\xi|$ by K', and the transmission coefficient

by Θ', one readily proves that

$$\Theta' = \frac{1}{1 + e^{-2K'}}, \qquad E \geqslant V_{max} \tag{21·42}$$

The reader will observe that these two expressions for the transmission coefficient join continuously at the intermediate case where $x_1 = x_2 = 0$, giving the common value $\frac{1}{2}$ for Θ when the energy level E just touches the top of the potential hill.

---

[1] In case the maximum value of $V(x) - E$ is very large (high hill), the Kramers connection formulas are applicable, but their unidirectional character leads to difficulty. It does not seem possible to set up a legitimate derivation of formulas for c and d in this way.

THE MATHEMATICAL THEORY OF COMPLETE SYSTEMS OF ORTHOGONAL FUNCTIONS

22. SCALAR PRODUCTS AND SYSTEMS OF ORTHOGONAL FUNCTIONS

22a. Expansion in a Series of Functions.

—It is well known that an arbitrary function of x, subject to certain continuity restrictions, can be developed into a Fourier's series

$$f(x) = \sum_{n=0}^{\infty}\left(a_n \cos \frac{nx}{c} + b_n \sin \frac{nx}{c}\right), \qquad (22\cdot1)$$

which is convergent and represents the function in the interval

$$0 \leqslant x \leqslant 2\pi c.$$

A similar development is possible in terms of the Hermitian orthogonal functions obtained in Sec. 20 as the eigenfunctions of the problem of the ideal linear oscillator in wave mechanics. In fact it is a property of a large class of similar boundary-value problems in one or more dimensions that they yield systems of eigenfunctions for which such a series development is both possible and convenient.

The importance of such a development will be evident from the following consideration. We assume that the expansion

$$f(x) = \sum_{n=0}^{\infty} a_n \psi_n(x) = \sum_{n=0}^{\infty} a_n c_n H_n(\sqrt{\alpha}x)e^{-\frac{1}{2}\alpha x^2} \qquad (22\cdot2)$$

is valid for any $f(x)$ which is quadratically integrable and piece-by-piece smooth[1] over the range $-\infty < x < +\infty$. Then with the aid of the eigenfunctions $\psi_n(x)$ we can at once determine the complete wave function $\Psi(x,t)$ for the harmonic-oscillator problem with arbitrary initial conditions. Suppose, for example, that $\Psi(x,0)$ is to have the form $f(x)$. The function

$$\Psi(x,t) = \sum_{n=0}^{\infty} a_n \psi_n(x) e^{-\frac{2\pi i E_n t}{h}} \qquad (22\cdot3)$$

---

[1] A function is said to be *"piece-by-piece" continuous* (*stückweise stetig*) in an interval K when the interval can be divided up into a finite number of subintervals such that the function is continuous in each and approaches a finite limit value as x approaches any of the subinterval boundaries. It is said to be *piece-by-piece smooth* (*stückweise glatt*) if it also has piece-by-piece continuous first derivatives in K.

reduces to $f(x)$ when $t = 0$, and if the series is term-by-term differentiable. it is a solution of the second Schrödinger equation appropriate to the problem, *viz.*,

$$\frac{\partial^2 \Psi}{\partial x^2} - \frac{4\pi^2 \mu k x^2}{h^2} \Psi + \frac{4\pi \mu i}{h} \frac{\partial \Psi}{\partial t} = 0. \tag{22·4}$$

Thus the series expansion enables us to pass from the special solution $\psi_n(x)$ to a solution of the most general type for the given mechanical system. It is also of great practical importance because it is the basis of most perturbation-theory computations.

Postponing for the present the discussion of the legitimacy of such an expansion as that given in Eq. (22·2) we note that the determination of the coefficients is greatly facilitated by the use of Eq. (20·10). To fix the value of any coefficient a_k we use the same procedure as in finding the coefficients of a Fourier's series. Multiply Eq. (22·2) by $\psi_k{}^*(x)$ and integrate term by term over the interval $-\infty < x < +\infty$. Equation (20·10) shows that all terms in this sum drop out except the one involving a_k. Since $\int_{-\infty}^{+\infty} \psi_n \psi_n{}^* dx = 1$,

$$a_k = \int_{-\infty}^{+\infty} f(x)\psi_k{}^*(x)dx. \tag{22·5}$$

The integral which forms the right-hand member of Eq. (22·5) is called the *scalar product* of the functions f and ψ and is indicated by the heavy round parenthesis symbol (f,ψ). The idea of the scalar product of two functions is of such fundamental importance in the mathematical theory of quantum mechanics that we pause here to review briefly the complex of ideas regarding the relations between vectors and sets of functions with which it is associated.

22b. Comparison of Properties of Vectors and Functions.—There is a very important and fundamental similarity between the properties of vectors in a space of n dimensions and the properties of quadratically integrable functions of a set of variables $x_1, \cdots, x_s$. This similarity lies at the basis of von Neumann's formulation of the problem of quantum mechanics in terms of transformations in Hilbert space.[1] Although a thoroughgoing treatment of Hilbert space and its relation to quantum mechanics lies outside the scope of this book, a brief discussion of the analogy between vectors and functions will be useful.

The basis for the similarity in question lies in the fact that the fundamental operations of ordinary vector algebra, with the exception of the formation of vector products, have parallels in important operations which may be applied to functions of one or more independent variables which are defined and quadratically integrable over a common

[1] *Cf.* JOHANN V. NEUMANN, *Göttinger Nachrichten*, 1927, p. 1; *M.G.Q.*; MARSHALL H. STONE, *Linear Transformations in Hilbert Space*, New York, 1932.

domain K in the space over which the independent variables range. (This latter space we shall hereafter designate as *configuration space*, whether the independent variables $x_1, \cdots x_s$ are the Cartesian positional coordinates of a dynamical system or not.) The first of the operations in question are the addition of vectors (or functions and the multiplication of a vector (or function) by a scalar or complex number. Both operations obey the ordinary rules of elementary algebra whether applied to vectors or to functions spread out over configuration space.

The third basic operation is the formation of the *scalar product* through which the idea of the *length* of a vector finds expression. The introduction of this operation converts the *affine geometry* of vectors into a metric geometry.[1] It will be useful briefly to recall the properties of the scalar product of two vectors before considering the definition of the scalar product of two functions.

22c. Scalar Products of Vectors.—In the case of two real three-dimensional vectors $\vec{A}, \vec{B}$ with components A_1, A_2, A_3, and B_1, B_2, B_3, respectively, the scalar product $\vec{A} \cdot \vec{B}$ is defined analytically by the equation[2]

$$\vec{A} \cdot \vec{B} = \sum_{k=1}^{3} A_k B_k. \tag{22·6}$$

This scalar product has the following important properties:

$$\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A}, \tag{22·7}$$

$$(a\vec{A}) \cdot \vec{B} = a(\vec{A} \cdot \vec{B}) = \vec{A} \cdot (a\vec{B}), \quad (a = \text{a scalar}) \tag{22·8}$$

$$(\vec{A} + \vec{A}') \cdot \vec{B} = \vec{A} \cdot \vec{B} + \vec{A}' \cdot \vec{B}, \tag{22·9}$$

$$|\vec{A} \cdot \vec{B}|^2 \leqq (\vec{A} \cdot \vec{A})(\vec{B} \cdot \vec{B}), \tag{22·10}$$

$$\vec{A} \cdot \vec{A} \geqq 0, \tag{22·11}$$

$$\vec{A} \cdot \vec{A} = 0 \text{ implies } \vec{A} = 0. \tag{22·12}$$

A complex vector in a space of n dimensions is simply an ordered set of n complex numbers. We define the Hermitian scalar product of two such vectors $\vec{A}, \vec{B}$ by the equation[2]

$$(\vec{A}, \vec{B}) = \sum_{k=1}^{n} A_k B_k^*. \tag{22·13}$$

---

[1] *Cf.* HERMANN WEYL, *Raum, Zeit, Materie*, 3d ed, Kap. 1, Berlin, 1920.

[2] We use a different notation for the scalar product of two three-dimensional vectors from that adopted for the scalar product of two n-dimensional vectors, or of two functions, because the two types of product sometimes occur in a close juxtaposition which might be confusing if the same symbol was used for both.

Here the complex conjugate of B_k is introduced instead of B_k itself to insure that when $\vec{B}$ is identified with $\vec{A}$ the scalar product $(\vec{A},\vec{A})$ shall be real and positive, or zero. The definition makes a slight formal change in the first two laws (22·7) and (22·8), which become

$$(\vec{A},\vec{B}) = (\vec{B},\vec{A})^* \tag{22·14}$$

and

$$(a\vec{A},\vec{B}) = a(\vec{A},\vec{B}) = (\vec{A},a^*\vec{B}), \tag{22·15}$$

respectively. Equations (22·9), (22·10), (22·11), and (22·12) hold as stated for complex vectors as well as for real vectors.

22d. Scalar Products of Quadratically Integrable Functions of m **Variables.**—Consider next a pair of complex functions, say f and g, of the real variables x_1, $x_2 \cdot \cdot \cdot x_m$ defined and piece-by-piece continuous in a certain domain K of configuration space. We symbolize the *scalar product* of f and g with respect to the domain K by (f,g) and define this quantity by the equation

$$(f,g) = \int_K fg^* dx_1 \cdot \cdot \cdot dx_m. \tag{22·16}$$

Usually the domain K is identical with the region of definition of f and g so that it need not be indicated explicitly on the scalar-product symbol.

The notation thus introduced is justified by the complete parallelism between the properties of the scalar product of two complex functions and the scalar product of two complex vectors, as will be seen by a comparison of the above equations with the following easily proved laws.

$$(f,g) = (g,f)^*, \tag{22·17}$$
$$(af,g) = a(f,g) = (f,a^*g), \tag{22·18}$$
$$(f + f',g) = (f,g) + (f',g), \tag{22·19}$$
$$|(f,g)|^2 \leqq (f,f)(g,g), \quad \text{(Schwarz's inequality)} \tag{22·20}$$
$$(f,f) \geqq 0, \tag{22·21}$$
$$(f,f) = 0 \text{ implies } f \equiv 0. \tag{22·22}$$

Equations (22·21) and (22·22) enable us to correlate (f,f) with the square of the length of a real vector, or with the square of the absolute value of a complex vector. It is sometimes called the *norm* of f [in symbols $(f,f) = Nf$] and becomes equal to unity when f is normalized according to the rule of Eq. (8·9). Similarly the square root of the norm of f is analogous to the length of a vector and may be called the *magnitude* of the function f. It is indicated by the symbol $\|f\|$, and has the basic property

$$\|f + g\| \leqq \|f\| + \|g\|. \tag{22·23}$$

$\|f - g\|$ is to be correlated with the distance from the end point of one

vector to the end point of another.   As a corollary on Eq. (22·22) we deduce that

$$\|f - g\| = 0 \text{ implies } f \equiv g. \tag{22·24}$$

In the quantum mechanics we have to do with functions whose range of definition K is infinite, and in which K is bounded by finite singular domains as well as by a surface at infinity.   Hence the scalar products are improper integrals, in general, and the following theorem is important.[1]

Theorem: The quadratic integrability of the functions f and g over a common domain K implies the absolute convergence of the scalar product (f,g) taken over K and the quadratic integrability of any linear combination of these functions over the same domain.

Proof: At any point in K,

$$|fg^*| = |f||g| \leqq \tfrac{1}{2}(|f|^2 + |g|^2).$$

Hence (f,g) is absolutely convergent.   Furthermore, at any point in K,

$$|f + g|^2 = |f|^2 + |g|^2 + (fg^* + f^*g). \tag{22·25}$$

As each of the terms on the right is integrable over K it follows that $|f + g|^2$ is integrable over K.   Introducing the fact that any multiple of a quadratically integrable function is quadratically integrable, we obtain at once the general theorem that $af + bg$ is quadratically integrable if a, b are constants and the integrals (f,f), (g,g) exist.

The above theorem justifies the last paragraph in Appendix C and leads to the following corollary.

Corollary: Any linear combination of a finite number of type A functions (for definition see p. 79) is a type A function.

The n-dimensional Function Space.—On the basis of the definitions of the operations of addition, multiplication by a complex number, and scalar multiplication, one can obtain an immediate one-to-one correspondence between the class of complex functions which are formed from linear combinations of any n linearly independent quadratically integrable functions $f_1, f_2 \cdots f_n$ of domain K and the class of complex vectors in a space of n dimensions.[2]   It will be convenient to establish this correspondence with the aid of the concept of orthogonality.

---

[1] *Cf.* J. von Neumann, *M.G.Q.*, p. 32.

[2] A set of n functions or vectors $f_1, f_2 \cdots f_n$ is said to be *linearly independent* if no one of them can be expressed as a linear combination of the others, or if a relation of the form

$$\sum_{i=1}^{n} a_i f_i = 0,$$

where the a's are complex numbers, can hold only if all the a's vanish.

Definition: If the scalar product of two vectors or of two functions having a common domain K is zero, we say that the vectors or functions are *orthogonal.*

Thus Eq. (20·10) asserts the mutual orthogonality of the different eigenfunctions of Eq. (20·1).

It follows as a corollary on the definition of orthogonality that any set of mutually orthogonal functions is necessarily linearly independent. Conversely, it can be proved that if we have given an arbitrary linearly independent set of quadratically integrable functions $f_1, f_2 \cdots f_n$, it is always possible to form a mutually orthogonal set from them by a suitable homogeneous linear transformation[1]

$$\sum_{k=1}^{n} \alpha_{ik} f_k = g_i. \qquad i = 1, 2, \cdots, n \qquad (22\cdot26)$$

Furthermore we can always choose this transformation so that the elements of the orthogonal set will be normalized according to the rule

$$(g_i, g_i) = 1. \qquad i = 1, 2, \cdots, n$$

In this case the functions g are said to form a *normal orthogonal* set, or, commonly, an *orthonormal* set.

Let us now employ such a normal orthogonal set to establish the previously mentioned correspondence between the class of functions which are expressible as linear combinations of the f's (call it class M) and the class of complex vectors in space of n dimensions. By means of Eqs. (22·26) we can express every function of class M as a linear combination of the g's. Thus, if ψ belongs to class M,

$$\psi = \sum_{i=1}^{n} a_i g_i \qquad (22\cdot27)$$

and the coefficients a_k are given by the formula

$$(\psi, g_k) = \sum_{i=1}^{n} a_i (g_i, g_k) = a_k. \qquad (22\cdot28)$$

Every such set of coefficients may be used to define a unique vector $(a_1, a_2 \cdots a_n)$ in a space of n dimensions, the individual coefficients being treated as components along a set of mutually orthogonal axes. Conversely, every vector in such a space yields a set of coefficients which determine a unique function which belongs to class M. Thus the one-to-one correspondence is established.

We can go much farther than this, however. Let V_ψ denote the vector which is correlated with the function ψ of class M. If we add

[1] *Cf.* Courant-Hilbert, *M.M.P.*, Kap. II, §1.

two such functions ψ and φ we obtain a new function, $\chi = \psi + \varphi$, which belongs to class M, and it is obvious that the vector V_χ of this new function is the sum of the vectors V_ψ and V_φ. Similarly, if we multiply ψ by an arbitrary complex number c, we obtain a new function of class M whose vector is cV_ψ. Finally, if we take the scalar product of two functions of class M we obtain a complex number which is exactly equal to the scalar product of the corresponding vectors. Thus

$$(\psi,\varphi) = \left(\sum_{i=1}^{n} a_i g_i, \sum_{i=1}^{n} b_i g_i \right) = \sum_{i=1}^{n} a_i b_i^*. \tag{22.29}$$

It follows as a corollary that the magnitude of the function ψ is equal to the absolute value of its vector. Consequently we have a basic structural similarity between the two classes of entities under consideration. They are said to be *isomorphous*.

The end points of all possible real vectors drawn from any fixed origin in a space of three dimensions fill up that space. Hence it is customary to apply the term *three-dimensional space* to the class of all such vectors. Similarly the class of complex vectors having n components is said to form an *n-dimensional space* and we can equally well apply this term to the class of all functions formed by linear combinations of n linearly independent functions. To distinguish such a class, or space, from a space composed of vectors we call it a *function space*.

22e. Spaces of Infinitely Many Dimensions.—More important varieties of function space are defined by the class of all type A solutions of a given Schrödinger equation and by the class, say L_{2K}, of all functions quadratically integrable over a common domain K in configuration space. These more general classes of functions can be correlated with the class of complex vectors of finite magnitude, or "length," and infinitely many dimensions. Such a vector is defined as an infinite sequence of complex numbers a_1, a_2, a_3, $\cdots$ such that $\sum_{i=1}^{\infty} a_i a_i^*$ is convergent. The operations of addition, multiplication by a number, and scalar multiplication are defined as for vectors in space of n dimensions.[1] The correlation of the elements of L_{2K} with the elements of this class of complex vectors is most easily formulated with the aid of an infinite orthonormal system of functions g_1, g_2, g_3, $\cdots$ in L_{2K} such that, if ψ is any other element of L_{2K}, the *completeness* relation

$$\lim_{n \to \infty} \left\| \psi - \sum_{i=1}^{n} a_i g_i \right\| = 0 \qquad a_i = (\psi,g_i) \tag{22.30}$$

---

[1] The infinitely many-dimensioned space defined by this class of vectors is called *Hilbert space*. It belongs along with L_{2K} to a class of classes all of which conform to the same set of postulates and which is called *abstract Hilbert space* (cf. references in footnote 1, p. 114).

is satisfied (*cf.* Sec. 25*b*, pp. 136 and 137). In physical applications this relation is usually equivalent to the validity of the series expansion

$$\psi = \sum_{i=1}^{\infty} a_i g_i, \qquad a_i = (\psi, g_i) \tag{22·31}$$

but not always (*cf.* Sec. 36*d*, p. 255). When such a complete orthonormal set of functions is given, we identify the scalar product a_i of ψ and g_i with the *i*th coordinate of the complex vector V_ψ and thereby establish the desired correlation. If ϕ is a second element of the function space L_{2K} with the components b_1, b_2, b_3, $\cdots$, it can be proved that the scalar product of ψ and ϕ is given by

$$(\psi, \phi) = \sum_{i=1}^{\infty} a_i b_i^*. \tag{22·32}$$

To demonstrate this proposition we note first of all that (22·30) is equivalent to

$$(\psi, \psi) = \sum_{i=1}^{\infty} a_i a_i^*.$$

Since this equation holds for any element of L_{2K}, it holds for the linear combination of functions $\chi = \psi + \lambda\phi$, where λ is an arbitrary parameter. Using the procedure of footnote 3, p. 36, we have

$$(\chi, \chi) = (\psi, \psi) + \lambda(\phi, \psi) + \lambda^*(\psi, \phi) + \lambda\lambda^*(\phi, \phi)$$

$$= \sum_{i=1}^{\infty} a_i a_i^* + \lambda \sum_{i=1}^{\infty} a_i^* b_i + \lambda^* \sum_{i=1}^{\infty} a_i b_i^* + \lambda\lambda^* \sum_{i=1}^{\infty} b_i b_i^*.$$

Since λ is arbitrary, the coefficients of λ and λ^* on the two sides of the equation must be equal independently. This proves (22·32).

The right-hand member of (22·32) forms the extrapolation to $n = \infty$ of the scalar product of two complex *n*-dimensional vectors. We conclude that we can identify the scalar product of two functions with the scalar product of their vector representatives in Hilbert space. This completes the definition of the scalar product of two functions and brings us back to the problem of series expansion with which we started the present section.

22f. Proof of Orthogonality of Eigenfunctions of the One-dimensional Anharmonic Oscillator Problem.—Although the mutual orthogonality of the functions g_i is not a necessary condition for the validity of an expansion such as $\psi = \sum_{i=1}^{\infty} a_i g_i$, it greatly facilitates the evaluation of the coefficients in such a series. Without this convenient property or an equivalent one (*cf.* Sec. 23), the determination of the coefficients would

be difficult, if not practically impossible. We have already seen [Eq. (20·10)] that the discrete eigenfunctions of the ideal linear oscillator problem of Sec. 20 are mutually orthogonal, and it will now be proved that the same property is shared by the type A eigenfunctions of the general one-dimensional oscillator problem of Eq. (18·2).

It follows from the above mentioned differential equation that if ψ_n and ψ_m are two discrete eigenfunctions having the eigenvalues E_n and E_m, respectively,

$$\frac{d}{dx}\left(\psi_n{}^*\frac{d\psi_m}{dx} - \psi_m\frac{d\psi_n{}^*}{dx}\right) = \psi_n{}^*\frac{d^2\psi_m}{dx^2} - \psi_m\frac{d^2\psi_n{}^*}{dx^2} = \frac{8\pi^2\mu}{h^2}(E_n - E_m)\psi_n{}^*\psi_m.$$

Integrating over all values of x, we obtain

$$\frac{8\pi^2\mu}{h^2}[E_n - E_m](\psi_m,\psi_n) = \left[\psi_n{}^*\psi_m{}' - \psi_m\psi_n{}^{*'}\right]_{-\infty}^{+\infty}. \qquad (22·33)$$

Since the eigenfunctions and their derivatives vanish at the boundary points $x = \pm\infty$, the right-hand member of Eq. (22·33) is zero. It follows that ψ_n and ψ_m are orthogonal unless $E_n = E_m$, in which case one of the functions is a multiple of the other.

The generalization of the above theorem for other one-dimensional problems will be given in the next section, and in Sec. 32d an extension to problems having many dimensions will be given.

23. SELF-ADJOINT OPERATORS AND EQUATIONS. THE STURM-LIOUVILLE PROBLEM

23a. Self-adjoint Differential Operators in One Dimension.—Since we shall encounter more general one-dimensional eigenvalue-eigenfunction problems than that of Sec. 18, it is important to investigate the conditions under which a problem of this type based on a second-order linear homogeneous differential equation will lead to a set of mutually orthogonal eigenfunctions. In other words, we need to develop an extension of the orthogonality theorem of Sec. 22 to equations of a more general character.

It is convenient to introduce the Hamiltonian operator

$$H = -\frac{h^2}{8\pi^2\mu}\frac{d^2}{dx^2} + V(x) \times \qquad (23·1)$$

and to rewrite Eq. (18·2) in the form

$$Hy - Ey = 0. \qquad (23·2)$$

The essential feature of the proof of the orthogonality of the eigenfunctions is evidently to be found in the identity

$$z(x)Hy(x) - y(x)Hz(x) \equiv \frac{h^2}{8\pi^2\mu}\frac{d}{dx}\left[y\frac{dz}{dx} - z\frac{dy}{dx}\right], \qquad (23·3)$$

which holds for arbitrary continuous and twice-differentiable functions $y(x)$ and $z(x)$. We therefore seek a generalization of Eq. (23·3).

Consider the second-order linear homogeneous differential equation

$$p_0(x)y'' + p_1(x)y' + p_2(x)y + \lambda\rho(x)y = 0. \qquad (23·4)$$

Here λ is a parameter which is to play the same role as E in Eqs. (18·2) and (23·2), while y' and y'' denote, respectively, the derivatives dy/dx and d^2y/dx^2. The coefficients $p_0(x)$, $p_1(x)$, $p_2(x)$, $\rho(x)$ are arbitrary but supposedly known functions of x, defined and continuous in the interior of a certain interval K, say $a < x < b$, of the x axis. Either a, or b, or both, may eventually become infinite. The functions $p_0(x)$, $p_1(x)$, $p_2(x)$ are permitted to take on complex values, but $\rho(x)$ and the independent variable x itself are restricted to real values. We further assume that $\rho(x)$ is of the same sign throughout the interval K. Introducing the symbol Λ for the linear operator,

$$\Lambda \equiv p_0(x)\frac{d^2}{dx^2} + p_1(x)\frac{d}{dx} + p_2(x) \times, \qquad (23·5)$$

we rewrite (23·4) in the symbolic form

$$\Lambda y + \lambda\rho y = 0. \qquad (23·6)$$

Let us now assume that the differential operator Λ satisfies a relation of the form

$$z^*(x)\Lambda y(x) - y(x)\Lambda^* z^*(x) \equiv \frac{dF}{dx}, \qquad (23·7)$$

where y and z are arbitrary functions and F denotes a bilinear form in the arguments y, y'; z^*, $z^{*\prime}$. In other words,

$$F \equiv a_0(x)y'z^{*\prime} + a_1(x)y'z^* + a_2(x)yz^{*\prime} + a_3(x)yz^*,$$

where $a_0(x)$, $a_1(x)$, $a_2(x)$, $a_3(x)$ are functions to be evaluated from $p_0(x)$, $p_1(x)$, $p_2(x)$. $z^*(x)$ is the complex conjugate of $z(x)$, and Λ^* is the operator which transforms z^* into $(\Lambda z)^*$. It is derivable from Λ by replacing $p_0(x)$, $p_1(x)$, $p_2(x)$ by their complex conjugates. A differential operator Λ for which a relation of the form (23·7) is valid is said to be *self-adjoint.*[1]

---

[1] In the case of a real differential operator the symbol Λ^* can be replaced by Λ and the definition of the self-adjoint property given above reduces to that stated in Courant-Hilbert, *M.M.P.* and other standard texts. In general, whether Λ conforms to (23·7) or not, there is a second-order differential operator $\Lambda^\dagger$ such that

$$z^*\Lambda y - y(\Lambda^\dagger z)^* \equiv \frac{dG}{dx},$$

where G again denotes a bilinear form in the arguments y, y'; z^*, $z^{*\prime}$ with coefficients to be evaluated from p_0, p_1, p_2. $\Lambda^\dagger$ is then said to be *adjoint* to Λ. However, a different definition of the term "adjoint" will be found useful at a later stage in the development of our theory (*cf.* Sec. 32, p. 203).

If Eq. (23·7) is to hold for all functions y and z^* the coefficients of each of the products yz^*, $yz^{*\prime}$, $y'z^*$, $\cdots$ on the two sides of the equation must be equal. Hence it follows[1] that Λ must have the special form

$$\Lambda = p_0 \frac{d^2}{dx^2} + (p_0' + if)\frac{d}{dx} + \left[\frac{if'}{2} + g\right] \times, \qquad (23\cdot5a)$$

where p_0, f, and g are real functions of x, and i denotes the square root of -1. The function F becomes

$$F = p_0(z^*y' - yz^{*\prime}) + ifyz^*. \qquad (23\cdot8)$$

Integrating Eq. (23.7) we obtain the Green's formula

$$\int_a^b (z^*\Lambda y - y\Lambda^*z^*)dx = F(b) - F(a). \qquad (23\cdot9)$$

The next step in the derivation of our orthogonality theorem is to assume that the self-adjoint equation (23·6) is to be solved subject to boundary conditions at $x = a$ and $x = b$. Identify y and z with two eigenfunctions y_1 and y_2, respectively. Then

$$\Lambda y = \Lambda y_1 = -\lambda_1 \rho y_1,$$
$$\Lambda^*z^* = \Lambda^*y_2^* = -\lambda_2^* \rho y_2^*.$$

Equation (23·9) becomes

$$\left.\begin{array}{l}(\lambda_2^* - \lambda_1)\int_a^b \rho y_1 y_2^* dx = \lim_{x\to b} \Phi(x) - \lim_{x\to a} \Phi(x), \\ \Phi(x) \equiv p_0(x)[y_2^*(x)y_1'(x) - y_1(x)y_2^{*\prime}(x)] + if(x)y_1(x)y_2^*(x).\end{array}\right\} \quad (23\cdot10)$$

If the boundary conditions insure that the right-hand member of this equation shall vanish, and if $\lambda_2^* \neq \lambda_1$, we see that the functions $\rho^{\frac12}y_1$ and $\rho^{\frac12}y_2$ are orthogonal in the region K. If y_1 and y_2 are identical, the integral cannot vanish because ρ does not change sign in K, and it follows that $\lambda^* = \lambda$. Thus, *the eigenvalues are all real*.

23b. Orthogonality with Respect to a Density Function ρ.—In the special case that ρ is constant, the eigenfunctions of Eqs. (23·6) and (23·5a) with suitable boundary conditions form an orthogonal system. If ρ is not constant, we may say that the functions are *orthogonal with respect*

---

[1] The equality of the coefficients yields the following relations:

$0 = a_0,$	(α)	$0 = a_1 + a_2,$	(δ)
$p_0 = a_1,$	(β)	$p_1 = a_1' + a_3,$	(ϵ)
$p_0^* = -a_2,$	(γ)	$p_1^* = -a_2' - a_3,$	(ζ)
		$p_2 - p_2^* = a_3'.$	(η)

Equations (β), (γ), (δ) show that p_0 must be real. The next pair of equations show that the real part of p_1 is equal to p_0' while the imaginary part is equal to a_3. Setting a_3 equal to if, we learn from the last equation that the imaginary part of p_2 is $f'/2$.

to the function ρ. If the functions $\rho^{\frac{1}{2}}y_n$ are quadratically integrable, it is convenient to normalize them so that

$$(\rho y_n, y_m) = \int^b \rho y_n y_m^* dx = \delta_{nm}. \tag{23.11}$$

The problem of expanding an arbitrary function in terms of a system of eigenfunctions having this modified type of orthogonality is no more difficult than in the case of direct orthogonality. If it is required to evaluate the coefficients in the series

$$f(x) = \sum_{n=0}^{\infty} a_n y_n(x), \tag{23.12}$$

we have only to multiply through by ρy_k^* and integrate in order to obtain

$$a_k = \int_a^b \rho(x)f(x)y_k^*(x)dx = (\rho f, y_k). \tag{23.13}$$

If $\rho^{\frac{1}{2}}f(x)$ is quadratically integrable over (a,b), or if $|f(x)|$ is bounded and $\rho y_k(x)$ is absolutely integrable, the coefficient a_k is sure to exist.

23c. The Sturm-Liouville Problem.—Usually the coefficients in Eq. (23.5a) are restricted to *real* values. Then the most general case in which the equation is self-adjoint is that in which it has the form[1]

$$\frac{d}{dx}\left(p\frac{dy}{dx}\right) - qy + \lambda\rho y = 0. \tag{23.14}$$

This is the Sturm-Liouville equation whose boundary-value problems are so fully discussed in various mathematical texts. It is assumed that all three of the coefficients $p(x)$, $q(x)$, $\rho(x)$ are continuous in the interior of the fundamental interval $a < x < b$ and that $p(x)$ and $q(x)$ are differentiable in the same region. To eliminate the possibility of a singular point in the fundamental interval we assume that $p(x)$ is positive throughout that interval. Finally we require that $\rho(x)$ shall be of one sign throughout the closed interval $a \leqslant x \leqslant b$.

---

[1] The notation for the coefficients of Eq. (23.14) has been altered from that of Eq. (23.5a) to conform to the usage of Courant-Hilbert.

The general second-order linear homogeneous differential equation (23.4) with real coefficients is thrown into the Sturm-Liouville form (23.14) if multiplied through by the function

$$f(x) = p_0^{-1}e^{\int p_1 p_0^{-1}dx}. \tag{a}$$

If the coefficients p_0, p_1, p_2 are continuous in the interval $a < x < b$, and if p_0 has no nodes between a and b, the equation will have no singular points in the same interval (*cf.* footnote 1, p. 79), and $f(x)$ will not vanish or become infinite there. We conclude that if Eq. (23.4) has no singular points in the interval $a < x < b$ it may be converted to Sturm-Liouville form without introducing such points.

Equation (18·2) is a simple special case of the Sturm-Liouville equation in which p and ρ are constants. A more general example arises in connection with the study of the vibrations of a stretched string with variable density and variable elastic modulus. The problem of solving Eq. (23·14) in the interval $a < x < b$ subject to linear homogeneous boundary conditions such as

$$\left.\begin{array}{l} \beta_1 y(a) - \beta_2 y'(a) = 0, \\ \gamma_1 y(b) + \gamma_2 y'(b) = 0, \end{array}\right\} \quad \beta_1, \beta_2, \gamma_1, \gamma_2 = \text{real constants} \quad (23\cdot15)$$

or

$$y(a) = y(b), \qquad p(a)y'(a) = p(b)y'(b), \qquad (23\cdot16)$$

is called the *Sturm-Liouville problem*. If β_2 and γ_2 are zero, we have the important special case in which $y(x)$ is required to vanish at the boundary points. These boundary conditions insure that if $y(x)$ is a solution of the problem, any constant multiplied into $y(x)$ will yield another solution for the same value of λ. Furthermore, the real and imaginary parts of every solution of a Sturm-Liouville problem are themselves solutions, so that the most general solution for a given value of λ can be obtained from the corresponding real solution, or solutions. There is then no appreciable loss of generality in restricting the discussion to real functions $y(x)$. The boundary conditions given above also have the property of reducing the right-hand member of Eq. (23·10) to zero and hence yield a set of eigenfunctions which are orthogonal with respect to $\rho(x)$ in the interval $a < x < b$.

The usual theory of the Sturm-Liouville problem assumes that the interval $a < x < b$ is finite and that the boundary points as well as all interior points are nonsingular.[1] In the wave mechanics, on the other hand, we meet a variety of one-dimensional eigenvalue problems based on equations of the form of (23·14) but involving singular end points and, frequently, infinite intervals $a < x < b$. These equations are usually obtained from multidimensional problems by the method of the separation of variables. Thus we are confronted with a need for an extension of the standard Sturm-Liouville theory covering the type of problem which arises in our physical discussions. Such an extension is sketched in this and the two succeeding sections.[2]

23d. Singular-point Boundary Conditions.—Every eigenvalue-eigenfunction problem based on a Sturm-Liouville equation and of physical

---

[1] *Cf.* footnote 1, p. 79.

[2] The singular end-point problems are of course an old story to mathematicians, but the writer has not been able to discover any comprehensive treatment of a sufficiently elementary character to meet the needs of physicists. The powerful and elaborate work of H. Weyl, *Math. Ann.* **68**, 220 (1910); *Nachrichten d. Kgl. Gesell. d. Wissenschaften zu Göttingen, Math.-phys. Klasse,* 1910, p. 1) covers most of the ground, but has for its primary purpose the study of the continuous spectrum and is not adapted to the requirements of beginners.

origin will make use of boundary conditions based on physical con-
siderations. In wave mechanics these *physical boundary conditions*
can be derived from the boundary and continuity conditions for type
A and type B functions given in Sec. 17, or from the more elaborate
conditions for physically admissible many-dimensional wave functions
formulated in Sec. 32b. Examples of such derivations are to be found in
Sec. 28. Owing to the variation in the physical boundary conditions
from problem to problem, and to a certain awkwardness in form, however,
it is convenient for the development of the general theory of the proper-
ties of eigenfunctions to replace them by a single, suitably chosen,
standard mathematical condition to which they are normally equivalent
in practice.

To this end the following *singular-point boundary condition* (s.p.b.c.)
for discrete eigenfunctions is proposed.[1] *The function $y(x)$ shall be said
to conform to the singular-point boundary condition for Eq.* (23·14) *at the
left-hand boundary point* $x = a$ *if* $\int_a^\xi \rho |y|^2 dx$ *exists when* $a < \xi < b$, *and
when a positive real number* ϵ *and a real number* m *exist, such that for
positive values* $x - a$ *in the neighborhood of* $x - a = 0$ *the functions*

$$\rho(x)|y|^2(x - a)^{-m}, \qquad \rho(x)|y'|^2(x - a)^{(m-\epsilon)}, \qquad (23\cdot17)$$

are bounded. The numbers ϵ *and* m *are to be independent of the parameter* λ
whose eigenvalues are sought. The corresponding condition for the right-
hand boundary point b is similar in form and need not be written down
explicitly. Either or both of the points a,b may be at infinity.

Although the above condition will seem somewhat formidable to the
average physicist, it is not difficult to apply. In fact we can ordinarily
set m equal to ϵ and so replace the requirement that the functions
(23·17) shall be bounded near $x = a$ by the simpler requirement that
$\rho(x)|y'|^2$ shall be bounded near $x = a$, while $\lim_{x \to a} \rho(x)|y|^2 = 0$. The
radial functions in the Dirac relativistic theory of the hydrogen atom
afford examples of the exceptional case of a function which satisfies
the more general condition but not the simplified one.

The singular-point boundary condition is said to be *equivalent* to the
physical boundary condition at $x = a$ in any particular case if an integral
curve of the given differential equation which conforms to the s.p.b.c.
at that point must also conform to the physical condition and *vice versa*.
The possibility of such equivalence in the case of two conditions not
really identical lies in the fact that the functions under consideration
are solutions of the differential equation and hence have only one or two

---

[1] *Cf.* E. C. KEMBLE, *Proc. Nat. Acad. Sci.* **19**, 710 (1933), where the condition is
given a slightly more general form. The article should be corrected by the insertion
on p. 711 of the additional stipulation that the function $g(x)$ is bounded in the neigh-
borhood of $x = a$.

types of behavior near $x = a$. This equivalence must be tested individually for each special problem by a study of the integral curves in the neighborhood of the singular boundary point under consideration. Methods for dealing with this question will be described later in this chapter. In the meantime it may be observed that the writer has found no case in which the equivalence of the physical boundary condition and the s.p.b.c. cannot be established without difficulty.

Let us next turn our attention to the general properties of the s.p.b.c. which make it useful.

a. If $u(x)$ conforms to the s.p.b.c. at $x = a$,

$$\lim_{x \to a} p(x)|u|^2 = 0.$$

b. If $u_1(x)$ and $u_2(x)$ conform to the s.p.b.c. at $x = a$,

$$\lim_{x \to a} \{p(x)|u_1||u_2'|\} = 0. \tag{23.18}$$

This relation follows directly from the fact that the square root of the product of $p(x)|u_1|^2(x - a)^{-m}$ and $p(x)|u_2'|^2(x - a)^{(m-\epsilon)}$ is bounded near $x = a$.

c. If $u_1(x)$ and $u_2(x)$ conform to the s.p.b.c. at $x = a$, any linear combination, say $w(x) \equiv \alpha u_1(x) + \beta u_2(x)$, will conform to the same condition. This is a consequence of the inequality

$$|w|^2 \leqslant 2[|\alpha u_1|^2 + |\beta u_2|^2]$$

(*cf.* p. 117). In this respect the singular-point condition shares one of the major properties of the homogeneous boundary conditions for regular boundary point problems [Eq. (23.15)].

d. Let $y_1(x)$ and $y_2(x)$ denote two linearly independent solutions of the Sturm-Liouville equation (23.14) for the same energy value. Then both of these cannot conform to the s.p.b.c. at either end point.

In order to verify the last of these properties it is convenient to introduce the function

$$W[y_1,y_2] \equiv p(x)[y_1(x)y_2'(x) - y_1'(x)y_2(x)].$$

It follows from the differential equation that this function is independent of x. If both y_1 and y_2 satisfy the s.p.b.c. at either boundary point it follows from the property **b** that $W[y_1,y_2]$ vanishes identically. As $p(x)$ is not zero, we conclude that

$$y_1(x)y_2'(x) = y_2(x)y_1'(x).$$

Integration of this equation shows that y_1 and y_2 are linearly dependent contrary to hypothesis. Hence both of them cannot satisfy the s.p.b.c.

An important corollary on the property **b** is that if y_1 and y_2 are eigenfunctions of a Sturm-Liouville problem with singular-point boundary

conditions and .the eigenvalues λ_1 and λ_2, respectively, the function $\Phi(x)$ of Eq. (23·10) vanishes at the end points. Hence the eigenvalues are real and the eigenfunctions are orthogonal with respect to the "density" function ρ. In this respect also the new boundary condition shares a fundamental property of the homogeneous boundary conditions of Eq. (23·15). We shall therefore classify eigenvalue-eigenfunction problems based on Eqs. (23·14), (23·17), and (23·18) along with those based on homogeneous conditions at regular boundary points as *Sturm-Liouville problems*. It follows from **d** that there cannot be two linearly independent eigenfunctions for such a problem with the same eigenvalue.

Solutions of the standard linear oscillator problem of Secs. 18 and 19 for Case II are readily seen to satisfy the singular-point boundary conditions and hence all general theorems proved for Sturm-Liouville problems with singular boundary points are applicable to the discrete spectra of linear oscillator problems.

*23e. Existence of Discrete Eigenvalues for Sturm-Liouville Problems with Singular End Points.—We have already discussed the existence and properties of discrete solutions of the standard linear-oscillator problem in Sec. 19. In order to deal with the general case of an arbitrary S.L. problem with singular end points, we postulate the existence of an interval G of λ values, say $\lambda' \leqslant \lambda \leqslant \lambda''$, such that for every λ in G there exists a pair of integral curves $u_\lambda(x)$, $v_\lambda(x)$ which conform to the s.p.b.c. at the boundary points a and b, respectively, and which do not have an infinite number of nodes in the neighborhood of these points. Proper normalization of these integral curves will then yield a pair of continuous functions of x and λ, say $u(x,\lambda)$, $v(x,\lambda)$, which for any given λ in G satisfy the conditions laid down for u_λ and v_λ.

If the function

$$W[u(x,\lambda),\ v(x,\lambda)] \equiv p(x)[u(x,\lambda)v'(x,\lambda) - v(x,\lambda)u'(x,\lambda)]$$

vanishes for any pair of values of x and λ, say ξ and $\bar{\lambda}$, it vanishes identically in x for $\lambda = \bar{\lambda}$ and the functions $u(x,\bar{\lambda})$, $v(x,\bar{\lambda})$ are linearly dependent. Hence *both* functions satisfy the s.p.b.c. at *both* of the end points, and $\bar{\lambda}$ is an eigenvalue with $u(x,\bar{\lambda})$ and $v(x,\bar{\lambda})$ as eigenfunctions. Every such eigenfunction has a finite number of nodes.

As λ increases, the spacing of the nodes of $u(x,\lambda)$ and $v(x,\lambda)$ decreases. The nodes of $u(x,\lambda)$ move to the left and those of $v(x,\lambda)$ to the right. We omit the proof of this proposition as it is readily carried through by standard methods (*cf.* footnote 1, p. 84), if we note that the function

$$W[u(x,\lambda_1),\ u(x,\lambda_2)] \equiv p(x)[u(x,\lambda_1)u'(x,\lambda_2) - u(x,\lambda_2)u'(x,\lambda_1)]$$

must vanish at $x = a$, while $W[v(x,\lambda_1),\ v(x,\lambda_2)]$ must vanish at $x = b$. Whenever a node of $u(x,\lambda)$ passes one of $v(x,\lambda)$, the function

$$W[u(x,\lambda),\ v(x,\lambda)]$$

must vanish at the common nodal point. Hence every such nodal crossing yields an eigenvalue of λ and an eigenfunction. There are no eigenvalues of λ for which the nodes of $u(x,\lambda)$ do not coincide with the nodes of $v(x,\lambda)$.

Let n' be the number of internal nodes of $u(x,\lambda')$ (lower limit of G), and let n'' be the number of internal nodes of $u(x,\lambda'')$ (upper limit of G). For simplicity we assume that λ' and λ'' are not eigenvalues and that n' is not zero. It follows at once from the above argument that *there exist $n'' - n'$ eigenvalues in the interval G, each having an eigenfunction with one more node than the eigenfunction of next lower eigenvalue.*

The case in which n' is zero requires special discussion since the eigenfunction of lowest eigenvalue has no internal nodes and cannot be obtained in general from a coincidence of the nodes of $u(x,\lambda)$ and $v(x,\lambda)$. It is convenient to change the dependent variable of Eq. (23·14) to

$$w(x) = p(x)^{1/2} y(x). \qquad (23 \cdot 19)$$

The differential equation now takes the form

$$w''(x) + \left[p^{-1}(\lambda \rho - q) - \frac{1}{2} p^{-1/2} \frac{d}{dx} \left(p^{-1/2} \frac{dp}{dx} \right) \right] w = 0. \qquad (23.20)$$

Let $w_1(x,\lambda)$ and $w_2(x,\lambda)$ denote the functions $p(x)^{1/2} u(x,\lambda)$ and $p(x)^{1/2} v(x,\lambda)$, respectively. Then w_1 vanishes at $x = a$, and w_2 at $x = b$. $W[u(x,\lambda), v(x,\lambda)]$ becomes $w_1 w_2' - w_2 w_1'$.

If $n' = 0$, and $w_1(b,\lambda') = 0$, the functions w_1 and w_2 are linearly dependent nodeless eigenfunctions of the eigenvalue λ'. If $w_1(b,\lambda') \neq 0$, the curves $w_1(x,\lambda')$ and $w_2(x,\lambda'')$ can be drawn in the upper half plane and will then intersect at some point $x = \xi$ where $w_1' < w_2'$. It follows that

$$W[u(x,\lambda'), v(x,\lambda')] = w_1(\xi,\lambda')[w_2'(\xi,\lambda') - w_1'(\xi,\lambda')] < 0.$$

Let λ_g denote a value of λ in the range G such that $w_1(x,\lambda_g)$ has a single interior node lying to the right of the corresponding node of $w_2(x,\lambda_g)$. Any eigenvalue of λ between λ' and λ_g must have a nodeless eigenfunction. But $w_1(x,\lambda_g)$ and $w_2(x,\lambda_g)$ must cross at some point ξ_g where $w_1' - w_2' < 0$. If we now vary λ between λ' and λ_g there must be some value of λ for which $w_1' - w_2'$ vanishes at the point where the curves cross. The function $W[u(x,\lambda), v(x,\lambda)]$ vanishes for this value of λ and the corresponding functions $u(x,\lambda)$, $v(x,\lambda)$ are linearly dependent eigenfunctions. Thus, if λ' and λ'' are not eigenvalues, the number of eigenvalues in G is always $n'' - n'$, even if n' is zero.

If $n'' > 0$ and the lower limit λ' of G is at $-\infty$, there is a minimum eigenvalue λ_0 with a nodeless eigenfunction.

To prove this theorem we have only to verify that for sufficiently large negative values of λ the curves for $w_1(x,\lambda)$ and $w_2(x,\lambda)$ when drawn in the upper half plane will intersect at a point where $w_1' > w_2'$. As a very large negative value of λ will make both curves convex to the axis in all but eventual infinitesimal neighborhoods of the end points, we may dispense with a complete analytic proof.

If the upper limit of G is $+\infty$, the sequence of ascending eigenvalues is denumerably infinite. If λ_n is the eigenvalue whose eigenfunction has n interior nodes, $\lim \lambda_n = \infty$.

In view of the above series of theorems the question of the existence and range of discrete eigenvalues for a Sturm-Liouville problem, subject to the singular-point boundary conditions, reduces to the problem of determining the range G (if any) in which the boundary conditions can be satisfied by real integral curves with a finite number of nodes. In case one or both of the boundary points is at infinity we can make the change of variables (23·19) and investigate the existence of solutions of the equation satisfying the boundary condition at infinity by the methods of Sec. 19. Singular boundary points located at finite points can frequently be tested by methods to be developed in Sec. 26.

In the case of a Sturm-Liouville problem with a finite fundamental region and regular end points subject to homogeneous boundary conditions, a parallel set of theorems can be proved. It is possible to show in particular that there is a denumerably infinite set of eigenvalues each having an eigenfunction with one more node than that of the next lower eigenvalue. There is a minimum eigenvalue with a nodeless eigenfunction.

24. REDUCTION OF EIGENVALUE PROBLEMS BASED ON SELF-ADJOINT DIFFERENTIAL EQUATIONS TO VARIATIONAL FORM[1]

Eigenvalue problems based on the differential equation

$$\Lambda y + \lambda \rho y = 0, \tag{24·1}$$

in which Λ denotes the self-adjoint operator of Eq. (23·5a), can be reduced to problems in the calculus of variations. This reduction is of great value both in the solution of concrete cases and in the development of general theorems. It can take a number of forms which we designate as A, B, C. We here summarize the results of calculations given in detail in Appendix E.

Variational Problem A.—Let J denote the integral

$$J[y,\lambda] = \int_a^b y^*[\Lambda y + \lambda \rho y]dx. \tag{24·2}$$

Let δy denote the first variation in $y(x)$ (*cf.* Appendix A) and let $F(x)$ be defined by

$$F(x) = p_0\left(y^*\frac{d}{dx}\delta y - \delta y\frac{dy^*}{dx}\right) + ify^*\delta y. \tag{24·3}$$

Let it be required to find a function $y(x)$ and a corresponding value for the parameter λ which satisfy the variational equation

$$\delta J = 0, \tag{24·4}$$

[1] *Cf.* Courant-Hilbert, *M.M.P.*, Chap. VI; Riemann-Weber, *D.P.*, vol. I, Chap. XX.

when the comparison functions are unrestricted except for the requirement of piece-by-piece continuous first and second derivatives, the condition that $J[y,\lambda]$ must exist, and for certain linear boundary conditions (α) which insure that[1]

$$\lim_{x \to a} F(x) - \lim_{x \to b} F(x) = 0. \tag{24·5}$$

The functions $y(x)$ which solve this problem are called *extremals*.

　Theorem a.[2]—*The required stationary values of J are all zero. They exist only for a set of eigenvalues of λ identical with the eigenvalues of Eq. (24·1) when solved subject to the same boundary conditions (α). The extremals of the variational problem and the eigenfunctions of the differential equation problem are identical.*

　Variational Problem B.—Let Q and N denote the integrals

$$Q[y] = -\int_a^b y^* \Lambda y \, dx; \qquad N[y] = \int_a^b \rho y y^* \, dx. \tag{24·6}$$

Stationary values of the ratio Q/N are required when the admissible comparison functions are subject to the same restriction as in the preceding problem and to the additional restriction that they shall yield convergent values of Q and N. (Since the continuous spectrum eigenfunctions are not quadratically integrable, the present method can yield only discrete eigenvalues.)

　Theorem b.[2]—*The stationary values of Q/N are identical with the eigenvalues of λ in the corresponding problem A and the extremals are identical with those of problem A.*

　Every Sturm-Liouville problem with a finite fundamental interval, regular end points, and homogeneous boundary conditions can be reduced to the form B. Sturm-Liouville problems involving infinite intervals and singular end points, subject to the singular-point boundary conditions can be reduced to the form B provided that the eigenfunctions yield convergent integrals N. This restriction is automatically fulfilled by the physically admissible class A wave functions of quantum mechanics.

　Variational Problem C.—Let it be required to find a function $y(x)$ which yields a stationary value of $Q[y]$ itself when the admissible comparison functions are subject to the restrictions of Problem B and also to the normalization condition $N[y] = 1$.

---

　[1] This requirement is a generalization of the boundary requirement for the orthogonality theorem of Sec. 23. Hence the latter requirement is always fulfilled if Eq. (24·5) holds for all admissible comparison functions. Evidently the condition that y shall vanish at the boundary points is normally sufficient. If Λ is real (Sturm-Liouville case), we may deduce Eq. (24·5) from the singular-point boundary condition [*cf.* Eq. (23·18)].

　[2] For proof *cf.* Appendix E.

Theorem c.[1]—*Every solution of Problems A and B when normalized by the introduction of a suitable constant factor yields a solution of Problem C. There are no other solutions of the latter problem. The stationary values of Q for the solutions of Problem C are the eigenvalues of* λ *for Problem A.*

Application to Sturm-Liouville Case.—An integration by parts yields the following alternative expression for the integral Q:

$$Q[y] = \int_a^b \left[p_0 y' y^{*\prime} - if y^* y' - \left(\frac{if'}{2} + g \right) yy^* \right] dx + p_0(a) y^*(a) y'(a)$$
$$- p_0(b) y^*(b) y'(b). \quad (24\cdot7)$$

This expression is particularly useful in the case of the Sturm-Liouville problems where the coefficient $f(x)$ vanishes identically. In applying Eq. (24·7) to this type of problem we revert to the notation of Eq. (23·14), replacing $p_0(x)$ by $p(x)$ and $g(x)$ by $-q(x)$. If the boundary conditions are of the singular-point type [*cf.* Eqs. (23·17) and (23·18)], we have

$$Q[y] = \int_a^b [p|y'|^2 + q|y|^2] dx \quad (24\cdot8)$$

for all admissible comparison functions.

25. COMPLETENESS OF SYSTEM OF DISCRETE EIGENFUNCTIONS OF A STURM-LIOUVILLE PROBLEM

*25a. The Eigenvalues as Absolute Minima.—A number of the most important properties of the Sturm-Liouville eigenfunctions and eigenvalues can be deduced from the application of the variational method to problems of the S. L. type. Courant and Hilbert have been leaders in the development of this mode of attack and readers are referred to their text for a more complete exposition of the results obtainable. We are interested here in the problem of the expansion of arbitrary functions in series of eigenfunctions of Sturm-Liouville problems and will follow the line of argument developed by Courant-Hilbert, indicating at the same time the way in which it can be extended to cover problems involving singular end points and infinite fundamental intervals.

Let us then consider a Sturm-Liouville problem having either regular end points or singular end points. If the boundary points are singular, we shall suppose that for each value of λ there is a one-parameter family of integral curves $\alpha u(x,\lambda)$ having a finite number of nodes which satisfies the left-hand boundary condition, and another such family $\beta v(x,\lambda)$ which satisfies the right-hand boundary condition, whereas integral curves not linearly dependent on $u(x,\lambda)$ or $v(x,\lambda)$ do not satisfy the corresponding boundary condition. The interval G of Sec. 23e will therefore extend from $\lambda = -\infty$ to $\lambda = +\infty$. In any such case it was proved in

[1] For proof *cf.* Appendix E.

Sec. 23e that there exists a nodeless eigenfunction $y_0(x)$ and a corresponding lowest eigenvalue λ_0. The eigenvalue-eigenfunction problem is reducible to the variational forms B and C.

It will now be proved that λ_0 is not only a stationary value of Q/N as required by Sec. 24, but that it is an absolute minimum of Q/N.

The transformed eigenfunction obtained from $y_0(x)$ by Eq. (23·19) will be denoted by $w_0(x)$. It is a solution of Eq. (23·20) which vanishes at the end points but has no interior nodes. If we express the integral Q of Eq. (24·8) in terms of $w(x)$, it takes the form

$$Q = \int_a^b \left[\left(w' - \frac{1}{2} \frac{p'}{p} w \right)^2 + \frac{q}{p} w^2 \right] dx, \tag{25·1}$$

while the normalizing integral becomes

$$N = \int_a^b \frac{\rho}{p} w^2 dx. \tag{25·2}$$

Here $w(x)$ is assumed to be real.

Our discussion is based on a theorem in the calculus of variations which states that if $y = U(x)$ is a nodeless extremal of the integral

$$D[y] = \int_{x_0}^{x_1} F(x,y,y') dx$$

subject to the boundary conditions

$$y(x_0) = \alpha, \qquad y(x_1) = \beta,$$

and if $F_{y'y'}(x,y,\bar{p}) > 0$ for every point x,y in the neighborhood of x, $U(x)$ and for every finite value $\bar{p}$, then U actually minimizes the integral D.[1] It is usually understood that the end points x_0,x_1 are regular, although, as we shall see, the theorem must hold for singular boundary points if it holds for regular ones.

We identify D with the integral

$$\mathcal{J}[y,\lambda] = \int_{a'}^{b'} (py'^2 + qy^2 - \lambda\rho y^2) dx = \int_{a'}^{b'} \left[\left(w' - \frac{1}{2} \frac{p'}{p} w \right)^2 + \frac{q}{p} w^2 - \frac{\lambda\rho w^2}{p} \right] dx, \tag{25·3}$$

restricting $p(x)$ to positive values as usual, and choosing a' and b' according to $a < a' < b' < b$. We identify the nodeless extremal $U(x)$ with $w_0(x) = p^{\frac{1}{2}}y_0(x)$, choosing the end-point conditions to suit, $i.e.$, writing them in the form

$$w(a') = \alpha \equiv w_0(a'), \qquad \psi(b') = \beta \equiv w_0(b'). \tag{25·4}$$

In the case of integrals of the quadratic type under consideration it can be shown that any minimum must be an absolute minimum[2] so that $\mathcal{J}[w_0,\lambda_0] \leqslant \mathcal{J}[w,\lambda_0]$ for all com-

---

[1] $Cf.$ O. BOLZA, $Lectures\ on\ the\ Calculus\ of\ Variations$, 22, p. 96, 1924. In stating the theorem we use the symbol $F_{y'y'}(x,y,\bar{p})$ for

$$\left. \frac{\partial^2 F(x,y,z)}{\partial z^2} \right|_{z=\bar{p}}.$$

[2] The second variation of $\mathcal{J}[w_0]$ is defined as $\delta^2\mathcal{J} = \frac{(\delta\alpha)^2}{2} \left\{ \frac{\partial^2}{\partial\alpha^2} \mathcal{J}[w_0 + \alpha\eta] \right\}_{\alpha=0}$. The first variation must vanish and $\delta^2\mathcal{J} \geqslant 0$ for every admissible variation $\eta(x)$ if w_0 minimizes $\mathcal{J}$. In the case of a quadratic integral of this type

$$\mathcal{J}[w_0 + \eta\delta\alpha] - \mathcal{J}[w_0] = \delta\mathcal{J} + \delta^2\mathcal{J} \geqslant 0.$$

As $w_0 + \eta\delta\alpha$ is an arbitrary admissible comparison function, we infer that the minimum is absolute.

parison functions w which conform to the boundary conditions (25·4). We designate the original problem of minimizing Q/N subject to the singular-point conditions and the problem of minimizing $\tilde{J}$ subject to (25·4) as problems X and Y, respectively.

Let $\tilde{Q}$ and $\tilde{N}$ be defined by

$$\tilde{Q}[w] = \int_{a'}^{b'} \left[\left(w' - \frac{1}{2}\frac{p'}{p}w \right)^2 + \frac{q}{p}w^2 \right]dx, \qquad \tilde{N}[w] = \int_{a'}^{b'} \frac{\rho}{p}w^2 dx \qquad (25\cdot5)$$

Then (*cf.* Sec. 24 and Appendix D),

$$\tilde{J}[w_0,\lambda_0] = \tilde{Q}[w_0] - \lambda_0\tilde{N}[w_0] = 0,$$

or

$$\lambda_0 = \frac{\tilde{Q}[w_0]}{\tilde{N}[w_0]}. \qquad (25\cdot6)$$

If $\eta(x)$ is an arbitrary continuous function of x, with piece-by-piece continuous first derivatives, which vanishes at the end points a',b',

$$\tilde{J}[w_0 + \eta,\lambda_0] - \tilde{J}[w_0,\lambda_0] = \tilde{Q}[w_0 + \eta] - \tilde{Q}[w_0] - \lambda_0(\tilde{N}[w_0 + \eta] - \tilde{N}[w_0])$$

$$= \tilde{N}[w_0 + \eta]\left(\frac{\tilde{Q}[w_0 + \eta]}{\tilde{N}[w_0 + \eta]} - \frac{\tilde{Q}[w_0]}{\tilde{N}[w_0]} \right) \geqslant 0.$$

Finally

$$\frac{\tilde{Q}[w_0 + \eta]}{\tilde{N}[w_0 + \eta]} \geqslant \frac{\tilde{Q}[w_0]}{\tilde{N}[w_0]} = \lambda_0, \qquad (25\cdot7)$$

so that λ_0 is an absolute minimum for $\tilde{Q}/\tilde{N}$.

If we were interested only in problems involving fixed regular end points we might have identified a' and b' with a and b from the beginning. Our theorem would then
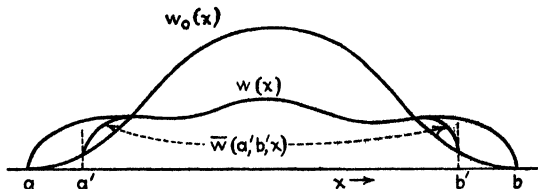
FIG. 8.

be proved. As it is we must resort to a limiting process to show that λ_0 is a minimum for Q/N as well as for $\tilde{Q}/\tilde{N}$.

Let $y(x)$ denote any admissible comparison function for the original Sturm-Liouville problem X in the complete interval $a < x < b$ and let $w = p^{1/2}y$. w does not automatically satisfy the boundary conditions of problem Y at a' and b', but it is easy to devise a function $\tilde{w}(a',b',x)$ which will satisfy them and which will be identical with $w(x)$ except in the immediate neighborhood of the points a' and b'. Furthermore, since $w(x)$ and $w_0(x)$ vanish at a and b, we can choose $\tilde{w}(a',b',x)$ so that it becomes equal to $w(x)$ in the limit when $a' = a$ and $b' = b$,[1] and so that

$$\lim_{a',b' \to a,b} \frac{\tilde{Q}[\tilde{w}]}{\tilde{N}[\tilde{w}]} = \frac{Q[w]}{N[w]}.$$

[1] Suppose, for example, that the graph of $w(x)$ lies above that for $w_0(x)$ in the neighborhood of the points $x = a$ and $x = b$. Then for points a' and b' not too far from a and b, respectively, the quantities $f(a')$ and $g(b')$ defined by

$$f(a') \equiv w(a') - w_0(a'), \qquad g(b') \equiv w(b') - w_0(b')$$

are positive. Under these circumstances we can choose $\tilde{w}(a',b',x)$ as follows:

In view of Eq. (25·7) $\tilde{Q}[\tilde{w}]/\tilde{N}[\tilde{w}]$ cannot approach a limit less than λ_0 and we conclude that

$$\frac{Q[w]}{N[w]} \geqslant \lambda_0,$$

which was to be proved.

It is possible to show that each of the higher eigenvalues is also a minimum value of Q/N for a suitably restricted class of admissible comparison functions. To be precise, let $y_n(x)$ denote an eigenfunction having n internal nodes. Then its eigenvalue λ_n is the minimum value of Q/N for all comparison functions which are orthogonal to ρy_0, ρy_1, ρy_2, $\cdots$, ρy_{n-1} and, in addition, conform to the usual continuity and boundary conditions.[1]

*25b. The Expansion of Arbitrary Functions in Terms of Eigenfunctions.—As we have already observed, the problem of the expansion of an arbitrary function in terms of an infinite series of mutually orthogonal functions is of fundamental importance in wave mechanics as in the older branches of mathematical physics. The first step toward the justification of such an expansion is to show that one can approximate the function as closely as desired in the sense of the theory of least squares by means of a suitable linear combination of members of the given orthogonal system.

To be specific, let $u_0(x)$, $u_1(x)$, $u_2(x)$, $\cdots$ denote an infinite series of functions spread out over the finite or infinite fundamental interval $a < x < b$ in which they are mutually orthogonal, quadratically integrable, and normalized. Let $f(x)$ denote an arbitrary function which is

$$\tilde{w}(a',b',x) = w(x) - f\left[\frac{f - (x - a')}{f + (x - a')}\right], \qquad a' < x < a' + f$$

$$\tilde{w}(a',b',x) = w(x), \qquad a' + f < x < b' - g$$

$$\tilde{w}(a',b',x) = w(x) - g\left[\frac{g + (x - b')}{g - (x - b')}\right]. \qquad b' - g < x < b'$$

The boundary conditions (25·4) are satisfied, and although $w(a',b',x)$ has a discontinuous slope at the points $x = a' + f(a')$ and $x = b' - g(b')$, it is an admissible comparison function for problem Y. The reader will readily verify that in this case, since $f(a) = g(b) = 0$, $\tilde{w}(a,b,x) = w(x)$ and

$$\lim_{a',b' \to a,b} \left(\frac{\tilde{Q}[\tilde{w}]}{\tilde{N}[\tilde{w}]}\right) = \frac{Q[w]}{N[w]}.$$

[1] In the case of a problem involving a finite interval and regular end points it is proved in Courant-Hilbert, *M.M.P.*, Kap. VI, §1, that if Q/N has a minimum for the specified type of comparison function, it must be λ_n. The existence of the minimum has been implicitly proved by Morse (*cf.* Marston Morse, "Sufficient Conditions in the Problem of Lagrange," *Am. J. Math.* **53**, 517 (1931); Kuen Sen Hu, Theorem 10.3, *Thesis*, University of Chicago, 1932, "Contributions to the Calculus of Variations," Chicago, 1933). The extension of the theorem to the case of problems with singular end points or infinite fundamental intervals can be carried through as in the case of the lowest eigenvalue.

piece-by-piece continuous and quadratically integrable in the same interval. To approximate the function f by means of the first n members of the series we form the expression

$$f_n(x) = \sum_{\nu=0}^{n} c_\nu u_\nu, \tag{25.8}$$

where the coefficients c are undetermined constants. The error of this approximation is

$$t_n = f - f_n = f - \sum_{\nu=0}^{n} c_\nu u_\nu, \tag{25.9}$$

and for the mean square of the error we may use the expression

$$W_n{}^2 = \int_a^b |t_n|^2 dx = \int_a^b f f^* dx - \sum_{\nu=0}^{n} [c_\nu(u_\nu, f) + c_\nu{}^*(f, u_\nu) - c_\nu c_\nu{}^*].$$

To minimize this function of the coefficients $c_0, c_1, \cdots, c_n$ we differentiate with respect to the real and imaginary parts of each of the c's and set the derivatives equal to zero. This yields the usual Fourier coefficient formula [*cf.* Eq. (22.5)]

$$c_\nu = \int_a^b f u_\nu{}^* dx \equiv (f, u_\nu). \qquad \nu = 0, 1, 2, \cdots, n \tag{25.10}$$

The expression for $W_n{}^2$ then reduces to

$$W_n{}^2 = (f, f) - \sum_{\nu=0}^{n} |c_\nu|^2,$$

from which we deduce Bessel's inequality

$$\sum_{\nu=0}^{\infty} |c_\nu|^2 \leqslant (f, f).$$

If the approximation f_n can be made as accurate as desired by increasing the number of terms, it is necessary that

$$\lim_{n \to \infty} W_n{}^2 = (f, f) - \sum_{\nu=0}^{\infty} |c_\nu|^2 = 0. \tag{25.11}$$

When this condition is satisfied for an arbitrary f we say that the system of functions $u_0, u_1, u_2, \cdots$ is *complete* [*cf.* Eq. (22.30)]. If the condition is satisfied, any function orthogonal to every member of the system must vanish identically, for in the case of such a function the c's are all zero and $\int_a^b |f|^2 dx = 0$. Clearly, if we remove any member of such a complete system the remaining system will be incomplete.

In the case of a system of functions which is *not directly orthogonal* but orthogonal with respect to an essentially positive weighting function $\rho(x)$, it is convenient to introduce a *modified definition of completeness*. Let us suppose, for example, that the series of functions u_0, u_1, u_2, $\cdots$ forms a normal orthogonal system in the sense that

$$(\rho u_i, u_j) = \delta_{ij}. \tag{25·12}$$

Let the accuracy of the approximation f_n be tested by the weighted mean square

$$W_n{}^2 = (\rho t_n, t_n) = (\rho f, f) - \sum_{\nu=0}^{n} [c_\nu(u_\nu, \rho f) + c_\nu{}^*(\rho f, u_\nu) - c_\nu c_\nu{}^*].$$

$W_n{}^2$ is a minimum when the coefficients are determined by the modified Fourier-coefficient formula [*cf.* Eq. (23·13)].

$$c_\nu = (\rho f, u_\nu). \tag{25·13}$$

If the approximation can be made as accurate as desired

$$\lim_{n \to \infty} W_n{}^2 = (\rho f, f) - \sum_{\nu=0}^{\infty} |c_\nu|^2 = 0, \tag{25·14}$$

and we again say that the series of functions $u_0, u_1, u_2, \cdots$ is *complete*.

Consider next the set of eigenfunctions $y_0, y_1, y_2, \cdots$ obtained by solving a Sturm-Liouville problem subject to boundary conditions which permit us to throw the problem into variational form. Let us further assume that λ_{n+1} is the minimum value of Q/N for functions which are orthogonal to ρy_0, ρy_1, ρy_2, $\cdots$, ρy_n and that (*cf.* Sec. 23)

$$\lim_{n \to \infty} \lambda_n = \infty. \tag{25·15}$$

Under these circumstances the conditionally orthogonal system of eigenfunctions is complete. The following proof is a modification of the one given in Courant-Hilbert, *M.M.P.*, Kap. VI, §3.

Although the definition of completeness can be extended to cover a system of functions which are not normalized, we assume that the eigenfunctions under discussion have been normalized in accordance with Eq. (25·12). We also make the (provisional) assumption that the arbitrary function $f(x)$ satisfies the same boundary and continuity restrictions as the admissible comparison functions for the variation problem. The error function t_n is defined as before by the equations

$$t_n = f - \sum_{\nu=0}^{n} c_\nu y_\nu. \qquad c_\nu = (\rho f, y_\nu) \tag{25·16}$$

It follows from the above definition that

$$(\rho y_i, t_n) = (t_n, \rho y_i)^* = 0. \qquad i = 0, 1, \cdots, n \tag{25·17}$$

Hence

$$(\Lambda y_i, t_n) = -\lambda_i(\rho y_i, t_n) = 0. \qquad i = 0, 1, \cdots, n \tag{25.18}$$

As Λ is self-adjoint, while f and t_n satisfy the boundary conditions,

$$(\Lambda y_i, t_n) - (y_i, \Lambda t_n) = p(b)[t_n^*(b)y_i'(b) - t_n^{*\prime}(b)y_i(b)]$$
$$- p(a)[t_n^*(a)y_i'(a) - t_n^{*\prime}(a)y_i(a)] = 0.$$

Thus

$$(\Lambda t_n, y_i) = (y_i, \Lambda t_n)^* = 0. \tag{25.19}$$

But since λ_{n+1} is the minimum value of Q/N for functions orthogonal to ρy_0, ρy_1, ρy_2, $\cdots$, ρy_n,

$$\lambda_{n+1} \leqslant \frac{Q[t_n]}{N[t_n]} = -\frac{(\Lambda t_n, t_n)}{(\rho t_n, t_n)} \tag{25.20}$$

It follows that as n becomes infinite either $(\Lambda t_n, t_n)$ must approach $-\infty$, or else $(\rho t_n, t_n)$ must approach zero. To ascertain the behavior of the numerator for large values of n we form

$$(\Lambda f, f) = (\Lambda t_n, t_n) - \sum_{\nu=0}^{n} |c_\nu|^2 \lambda_\nu + \sum_{\nu=0}^{n} [c_\nu^*(\Lambda t_n, y_\nu) + c_\nu(\Lambda y_\nu, t_n)].$$

The second summation vanishes due to Eqs. (25.18) and (25.19). Hence $(\Lambda t_n, t_n)$ is the sum of a term which is independent of n and of another which increases with n when n is large. We conclude that

$$\lim_{n \to \infty} (\Lambda t_n, t_n) \neq -\infty,$$

and that consequently

$$\lim_{n \to \infty} (\rho t_n, t_n) = \lim_{n \to \infty} w_n^2 = 0. \tag{25.21}$$

This proves the completeness of the system of eigenfunctions insofar as functions f satisfying the boundary conditions are concerned. For the removal of this restriction and the extension of Eq. (25.14) to all piece-by-piece continuous quadratically integrable functions see Courant-Hilbert, Kap. VI, §3.

It follows from the definition of completeness that the series

$$\sum_{\nu=0}^{\infty} c_\nu y_\nu, \qquad c_\nu = (\rho f, y_\nu)$$

converges *on the average* in the interval a,b. Since the mean square of the error function for an n term approximation approaches zero in the limit, we say that the completeness of the sequence of functions y_ν implies that the series $\sum_{\nu=0}^{\infty} c_\nu y_\nu(x)$ converges "in the mean square" to the limit function $f(x)$. An investigation by Weyl cited in Sec. 30 indicates that if $f(x)$ satisfies all the restrictions placed on the comparison functions of the variation problem, the series will not only converge in mean square,

but will converge uniformly at every point to the value $f(x)$.[1]   Hence we shall assume for such functions that

$$f(x) = \sum_{\nu=0}^{\infty} (\rho f, y_\nu) y_\nu(x). \tag{25.22}$$

If the functions $y_\nu(x)$ are not normalized, the expansion becomes

$$f(x) = \sum_{\nu=0}^{\infty} \left[\frac{\int_a^b \rho f y_\nu{}^* dx}{\int_a^b \rho |y_\nu|^2 dx} \right] y_\nu(x). \tag{25.23}$$

Our proof of the completeness theorem for the eigenfunctions of a Sturm-Liouville problem is limited to the case in which the discrete eigenvalue spectrum extends to $+\infty$.   In Sec. 30 a modified statement of the theorem is formulated which applies when there is a continuous spectrum of eigenvalues extending to infinity and replacing or supplementing the discrete spectrum.   In such cases, also, there is a convenient expansion theorem applicable to physically admissible functions and containing an integral over the continuous spectrum of B type eigenfunctions in addition to the sum over the discrete functions.

In mentioning these expansions, however, it should be remarked that in the last analysis they are less general than the completeness relation itself—(25.11) holds for an arbitrary quadratically integrable $f(x)$—and so less necessary.   We shall return to this point in Secs. 30 and 36.

As a corollary on the completeness theorem we have the following formula for the scalar product of two functions $f(x)$, $g(x)$ which are piece-by-piece continuous and quadratically integrable.

$$(\rho f, g) = \sum_{\nu=0}^{\infty} c_\nu d_\nu{}^*, \qquad d_\nu = (\rho g, y_\nu) \tag{25.24}$$

[cf. Eq. (22.32) ff.].

[1] Courant-Hilbert give a proof of (25.22) that covers all cases in which the functions $y_\nu(x)$ are bounded and the interval $a \leqslant x \leqslant b$ is finite (cf. Courant-Hilbert, M.M.P., pp. 370–371)

THE DISCRETE ENERGY SPECTRUM OF THE TWO-PARTICLE CENTRAL-FIELD PROBLEM

26. THE BEHAVIOR OF SOLUTIONS OF AN ORDINARY SECOND-ORDER DIFFERENTIAL EQUATION NEAR A SINGULAR POINT

If the end points of the fundamental interval $a < x < b$ of a Sturm-Liouville problem are singular points of the differential equation,[1] it is of the greatest importance, both for proving that the problem has solutions and for finding the solutions, to know some of the results of the general theory of the behavior of solutions of second-order linear homogeneous differential equations in the neighborhood of such points.[2] In stating the results of the theory we restrict ourselves *a priori* to the case in which the coefficients in Eqs. (23·4) and (24·1) are single-valued and analytic except at certain isolated, irremovable singular points. (As the average reader knows, a function $y(x)$ is said to be *analytic* at a point $x = a$ provided that it can be developed in a Taylor's series,

$\sum\limits_{n=0}^{\infty} a_n(x - a)^n$, which converges in the neighborhood of that point.

The function $y(x)$ is said to have an *isolated singular point* at $x = x'$ if it is analytic at all points in the neighborhood of x' but is not analytic at x' itself.)

Let us now divide Eq. (23·4) through by $p_0(x)$, rewriting it in the standard form

$$\frac{d^2y}{dx^2} + p_1(x)\frac{dy}{dx} + p_2(x)y = 0. \qquad (26·1)$$

It can be proved that when the equation is written·in this form *the solutions are all analytic wherever the coefficients are analytic* [*cf., e.g.,* Bieberbach, *loc. cit.,* (p. 207)]. But since the zeros of $p_0(x)$ generate poles in the new coefficients p_1 and p_2, it will be seen that the singular points of the differential equation, as defined in Sec. 17, are now simply singular points for one or the other of the coefficients. In the neighbor-

---

[1] *Cf.* footnote 1, p. 79.

[2] *Cf.* E. L. INCE, *Ordinary Differential Equations,* pp. 160–168, 356–370, London and New York, 1927; L. BIEBERBACH, *Differentialgleichungen,* 3d ed., pp. 206–219, Berlin, 1930; RIEMANN-WEBER, *D.P.,* pp. 248–254.

hood of an ordinary (*i.e.*, nonsingular) point of the equation, say $x = a$, every solution is expansible in a convergent series

$$y = \sum_{\nu=0}^{\infty} \alpha_\nu (x - a)^\nu. \tag{26.2}$$

Solutions can also be analytic at the singular points of the equation in some cases, but, as previously noted, have a tendency to become discontinuous at such points.

When the development (26·2) does fail, the next possibility in order of simplicity is an expansion of the form

$$y = (x - a)^r \sum_{\nu=0}^{\infty} \alpha_\nu (x - a)^\nu. \qquad \alpha_0 \neq 0 \tag{26.3}$$

It can be proved that at least one solution exists in the neighborhood of $x = a$ for which the expansion (26·3) is possible, provided that the point in question is a pole for one, or both, of the coefficients $p_1(x)$ and $p_2(x)$ of order not greater than one for p_1 and two for p_2. In other words, a solution of the form (26·3) exists, if Eq. (26·1) can be written in the form

$$\frac{d^2 y}{dx^2} + \frac{P_1(x - a)}{x - a} \frac{dy}{dx} + \frac{P_2(x - a)}{(x - a)^2} y = 0, \tag{26.4}$$

where P_1 and P_2 are analytic at $x = a$ and hence expansible in power series about that point. When this condition is fulfilled the point $x = a$ is said to be a *regular singular point* ("*Stelle der Bestimmtheit*"). Otherwise it is an irregular singular point.

To obtain the solution in the neighborhood of a regular singular point, one may expand P_1 and P_2 in the series

$$P_1 = \sum_{\nu=0}^{\infty} \beta_\nu (x - a)^\nu, \qquad P_2 = \sum_{\nu=0}^{\infty} \gamma_\nu (x - a)^\nu,$$

insert the right-hand member of Eq. (26·3) for y into Eq. (26·4), and collect terms involving like powers of $(x - a)$. The sum of the coefficients of each such set of terms should vanish. The series of equations obtained from this condition determine both the exponent r and the relative values of the successive coefficients α_ν. From the terms of lowest power one obtains the equation for r, *viz.*,

$$r(r - 1) + r\beta_0 + \gamma_0 = 0. \tag{26.5}$$

This is known as the *fundamental equation* or *indicial equation* for the singular point. If the roots of this equation are distinct, and do not differ by an integer, there are two linearly independent solutions of the equation having the form (26·3), one for each root.[1]

---

[1] Any solution of a second-order linear homogeneous differential equation can be expressed as a linear combination of two linearly independent solutions. A pair of

On the other hand, if the roots of the fundamental equation are equal, or differ by an integer, the root r_1 having the larger real part[1] is to be used to give a solution, say y_1, of the form (26·3) and the second linearly independent solution can be proved to have the form

$$y_2 = y_1 A \log (x - a) + (x - a)^{r_2} \varphi(x - a). \qquad (26·6)$$

Here $\varphi(x - a)$ is analytic at $x = a$ and does not vanish there. The exact form of φ may be determined like that of y_1 by expansion in series and the method of undetermined coefficients. A is a constant which is not zero if $r_1 = r_2$, but *can* vanish when $r_1 - r_2$ is any other integer.

When the point a is an *irregular singular point* the solution in its neighborhood is more difficult to handle, though convergent series involving an infinite number of negative as well as of positive powers of x are known to exist. Asymptotic but ultimately divergent series involving only a finite number of negative powers of x also yield useful representations of the solutions of differential equations in the neighborhood of such points.[2]

The Irregular Singular Point at Infinity.—If the transformation

$$x = \frac{1}{z}, \qquad y\left(\frac{1}{z}\right) \equiv u(z) \qquad (26·7)$$

yields an equation for $u(z)$ with a singular point at the origin, we say that the equation for $y(x)$ has a singular point at infinity. If $u(z)$ has a regular singular point at the origin, $y(x)$ is said to have a regular singular point at infinity, and the behavior of $y(x)$ for large values of x is determined by the behavior of $u(z)$ for small values of z.

Consider, for example, the linear oscillator problem of Sec. 18. After applying the transformation (26·7) to (18·2), we obtain

$$u'' + \frac{2u'}{z} + \frac{8\pi^2\mu}{h^2}\left[E - V\left(\frac{1}{z}\right) \right]z^{-4}u = 0 \qquad (26·8)$$

as the differential equation for $u(z) = y(1/z)$. On account of the term in $u'(z)$ the origin is necessarily a singular point. The condition for a regular singular point is that $[E - V(1/z)]z^{-2}$ shall be analytic at $z = 0$, and cannot possibly be satisfied for all values of E.

Fortunately we have already derived in Sec. 19 and Appendix C the essential information regarding the behavior of solutions of the linear oscillator equation at

---

linearly independent solutions is called a *fundamental system*. From any one such system an infinity of others can always be formed by making linear combinations. The problem of solving a differential equation in the neighborhood of a point $x = a$ may be regarded as finished when two such linearly independent particular solutions are found.

[1] This rule covers the general case where β_0 and γ_0 may be complex. If they are real, the roots of Eq. (26·5) are conjugate complex quantities, or else real. In the former case, or in the case of equal roots, either one may be used.

[2] *Cf.* L. Schlesinger, *Einführung in die Theorie der gewöhnlichen Differentialgleichungen*, 3rd ed., Kap. 8, Berlin, 1922; L. Bieberbach, *Differentialgleichungen*, 3d ed.. Abschnitt 2, Kap. 4, §10, Berlin, 1930.

infinity.   Thus we know that if $V(x)$ approaches a positive limit greater than E when x approaches infinity, or if $V(x)$ becomes positively infinite as any finite power of x, there is one and only one integral curve for a given value of E which satisfies the singular-point boundary condition at $x = \infty$.   On the other hand, if $[V(x) - E]$ approaches a finite negative limit as x becomes infinite, the curves all continue to oscillate as x increases, and none of them satisfies the singular-point condition, although all remain finite and conform to the continuity condition of Sec. 17.   We shall return to the discussion of the properties of this latter class of integral curves in Sec. 30.

27. THE LEGENDRE POLYNOMIALS

27a. General Properties of the Legendre Equation.—As an example

of the application of the mathematical theory developed in Secs. 22 to 26 let us consider the special Sturm-Liouville differential equation

$$\frac{d}{dx}\left[(1 - x^2)\frac{dy}{dx}\right] + \lambda y = 0 \qquad (27\cdot1)$$

in the interval $-1 \leqslant x \leqslant 1$.   The equation has singular points at $x = \pm 1, \pm \infty$.   In order to examine the point $x = +1$ we throw the equation into the standard form (26·4) and find that

$$P_1(x - 1) = \frac{2x}{(x + 1)}; \qquad P_2(x - 1) = \frac{(1 - x)\lambda}{(x + 1)}.$$

Both of these functions are analytic at $x = 1$.   Hence this is a regular singular point.

The constant terms in the expansions of P_1 and P_2 about $x = 1$ are

$$\beta_0 = 1, \qquad \gamma_0 = 0,$$

respectively.   The fundamental equation reduces to

$$r(r - 1) + r = r^2 = 0. \qquad (27\cdot2)$$

As the roots are both zero, the fundamental system of solutions y_1, y_2 in the neighborhood of $x = 1$ reduces to the form

$$y_1 = \sum_{\nu = 0}^{\infty} \alpha_\nu(1 - x)^\nu; \qquad (27\cdot3)$$

$$y_2 = y_1 A \log (1 - x) + \sum_{\nu = 0}^{\infty} \delta_\nu(1 - x)^\nu. \qquad (27\cdot4)$$

Evidently y_1 is analytic at $x = 1$.   Furthermore, it cannot vanish there since substitution of the power series (27·3) into the differential equation yields the recurrence formula

$$\alpha_{\nu+1} = \alpha_\nu\left[\frac{-\lambda + \nu(\nu + 1)}{2(\nu + 1)^2}\right], \qquad (27\cdot5)$$

from which it follows that if y_1 vanishes at $x = 1$ (*i.e.*, $\alpha_0 = 0$), it must vanish identically.

y_1 conforms to the singular-point boundary condition (23·17) at $x = 1$ for all values of λ with the constant m set equal to zero. It has a finite number of nodes in the neighborhood of the boundary point in question.

The symmetry of the differential equation with respect to the transformation $x \rightarrow -x$ shows that the point $x = -1$ is also a regular singular point with the same fundamental equation (27·2). We conclude that for all values of λ there are integral curves which satisfy the singular-point condition at $x = -1$ and have a finite number of nodes in the neighborhood of that point.

Thus there is an interval G extending from $\lambda = -\infty$ to $\lambda = +\infty$ and conforming to the requirements of the initial existence theorem of Sec. 23e. The theorems of Sec. 23e show further that there is a minimum eigenvalue with a nodeless eigenfunction and that the infinite sequence of ascending eigenvalues satisfies the condition

$$\lim_{n \to \infty} \lambda_n = \infty.$$

It follows from Sec. 23b that the eigenfunctions are mutually orthogonal, and from Sec. 24 that the eigenvalues are the stationary values of Q/N, where

$$Q = \int_{-1}^{+1}(1 - x^2)|y'|^2 dx, \qquad N = \int_{-1}^{+1}|y|^2 dx$$

and the boundary conditions are of the singular-point type. Since there is a nodeless eigenfunction, Q/N must have a minimum value. Direct examination shows that Q and N are essentially positive, so that

$$\lambda_0 = \left(\frac{Q}{N}\right)_{\min} > 0.$$

Finally the eigenfunctions form a *complete* orthogonal system.

27b. Explicit Determination of Eigenvalues and Eigenfunctions.—
The eigenvalues of λ can be derived from the series (27·3). Let u_ν denote the νth term of this series. Since

$$\lim_{\nu \to \infty}\left(\frac{u_{\nu+1}}{u_\nu}\right) = \lim_{\nu \to \infty}\left[\left(\frac{\lambda - \nu(\nu + 1)}{(\nu + 1)^2}\right)\frac{(x - 1)}{2}\right] = \frac{1 - x}{2},$$

the boundary point $x = -1$ is the limit of convergence unless the series terminates with the nth term due to the circumstance that

$$\lambda = n(n + 1). \qquad n = 0, 1, 2, 3, 4, \cdots \qquad (27\cdot6)$$

Let us assume that the series does not so terminate. Then for sufficiently large values of ν, say $\nu > h$, we can neglect λ in comparison with $\nu(\nu + 1)$.

It follows from (27·5) that to this degree of approximation

$$2^h h \alpha_h = 2^{h+1}(h+1)\alpha_{h+1} = \cdots = 2^\nu \nu \alpha_\nu = \cdots$$

or

$$\alpha_\nu = \frac{\text{constant}}{2^\nu \nu}.$$

Thus the higher members of the series approximate to the corresponding members of the expansion of $\log \left(1 + \dfrac{x-1}{2}\right)^{-1}$ in powers of $(1-x)/2$, and the sum of the series must approach infinity, as x approaches the point -1 in such a way that the singular-point boundary condition is not satisfied.[1] We conclude that the values of λ given by Eq. (27·6) are the only ones which permit solutions of our problem and are hence the desired eigenvalues.

The eigenfunctions are polynomials which can now be derived from Eqs. (27·3) and (27·5). A more convenient procedure, however, is to use a power-series expansion about the mid-point of the fundamental interval:

$$y = \sum_\nu b_\nu x^\nu.$$

If we substitute this series into the differential equation (27·1), give λ one of its eigenvalues, and equate the coefficient of each power of x to zero, we obtain the recurrence formula

$$b_{\nu+2} = \frac{\nu(\nu+1) - n(n+1)}{(\nu+1)(\nu+2)} b_\nu.$$

This shows that the odd and even powers of x form independent series as in the linear-oscillator problem of Sec. 20. If n is even, the even series terminates; if n is odd, the odd series terminates. The eigenfunctions are evidently the polynomials obtained by suppressing the independent infinite series. Apart from a constant normalizing factor they are identical with the well-known Legendre polynomials, or zonal spherical harmonics. The latter are defined by the formulas

$$P_0(x) = 1; \qquad P_n(x) = \frac{1}{2^n n!} \frac{d^n(x^2-1)^n}{dx^n}. \qquad n = 1, 2, \cdots \qquad (27\cdot7)$$

A summary of the most important properties of the Legendre polynomials is given in Appendix F, together with explicit formulas for the early members of the series.

[1] *Cf.* Courant-Hilbert, *M.M.P.*, p. 281.

28. THE ENERGY LEVELS OF THE TWO-PARTICLE PROBLEM

28a. The Wave Equation.—The problem of two interacting particles is the fundamental eigenvalue problem of the Schrödinger theory. We assume that the potential energy V depends only on the distance r between the two particles. Then, as in Sec. 15, it is convenient to introduce the coordinates of the center of gravity and the components of r as independent variables. Let x_1, y_1, z_1, x_2, y_2, z_2 be the absolute coordinates of the two particles and let x, y, z be defined by

$$x = x_2 - x_1, \qquad y = y_2 - y_1, \qquad z = z_2 - z_1. \qquad [Cf.\ \text{Eqs. (15·14)}]$$

From the general equation (17·1) we may deduce the wave equation for the internal motion of the system when there is a definite internal energy E. It is

$$\frac{1}{\mu}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi + \frac{8\pi^2}{h^2}(E - V)\psi = 0, \qquad (28\cdot1)$$

where μ is defined by

$$\frac{1}{\mu} = \frac{M}{\mu_1\mu_2} = \frac{1}{\mu_1} + \frac{1}{\mu_2}. \qquad (28\cdot2)$$

This is also the wave equation for a *single* particle of mass μ moving under the influence of a *fixed* center of force at the origin of the x,y,z system of coordinates.

28b. Separation of the Variables.—To solve Eq. (28·1), introduce spherical coordinates r, θ, φ.[1] Direct transformation then yields

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2} + \kappa(E - V)\psi = 0. \qquad (28\cdot3)$$

Here κ denotes the quantity $(8\pi^2\mu)/h^2$ as usual.

The variables can now be separated as in the derivation of Eqs. (15·17) and (15·18). We seek particular solutions of the form

$$\psi = R(r)Y(\theta,\varphi). \qquad (28\cdot4)$$

Inserting this expression into Eq. (28·3) and rearranging, we obtain

$$\frac{1}{R}\left\{\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \kappa r^2\left[E - V(r)\right]\right\} = -\frac{1}{Y}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2}\right].$$

One side of this equation is a function of r only, while the other side depends upon the polar angle θ and the azimuthal angle φ, but not on r. Hence the two sides must have a common constant value which we shall

---

[1] θ is the angle between the radius vector and the z axis while φ is the azimuthal angle between the xz plane and a plane through the radius vector and the z axis. (*Cf.* Courant-Hilbert, *M.M.P.*, p. 194, or Riemann-Weber, *D.P.*, Vol. I, p. 76.)

call λ, and the equation is equivalent to the pair of simultaneous equations

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \kappa\left[E - V(r) - \frac{\lambda}{\kappa r^2}\right]R = 0; \qquad (28\cdot5)$$

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2} + \lambda Y = 0. \qquad (28\cdot6)$$

Equation (28·6) plays an important part in potential theory[1] where it must be solved in such a way that Y is single-valued and continuous over the entire sphere. Such solutions are called *surface harmonics*. The same requirements apply to our problem.

The variables in Eq. (28·6) may be separated in turn if we ask for special solutions of the form

$$Y = \Theta(\theta)\Phi(\varphi). \qquad (28\cdot7)$$

The equation then breaks up into the pair of equations

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) - \frac{\beta\Theta}{\sin^2\theta} + \lambda\Theta = 0; \qquad (28\cdot8)$$

$$\frac{d^2\Phi}{d\varphi^2} + \beta\Phi = 0. \qquad (28\cdot9)$$

28c. The Azimuthal Factor of the Wave Function.—If $\psi(x,y,z)$ is to be a single-valued function of the Cartesian coordinates, it follows that $\Phi(\varphi + 2\pi) = \Phi(\varphi)$. In view of this condition the only admissible solutions of Eq. (28·9) are the familiar harmonic functions

$$\Phi = A\begin{cases}\sin m\varphi, \\ \cos m\varphi, \qquad m = 0, \pm1, \pm2, \cdots \\ e^{im\varphi}.\end{cases} \qquad (28\cdot10)$$

Thus the eigenvalues of β are $0^2, 1^2, 2^2, 3^2, \cdots$, In this case, due to the special boundary condition, there are two linearly independent solutions for each eigenvalue—a phenomenon called "degeneracy."

The normalization condition of Sec. 8 requires that the integral of $\psi\psi^*$ over all space shall be unity. In spherical coordinates this means that

$$\int_0^\infty RR^*r^2dr\int_0^\pi \Theta\Theta^* \sin\theta\, d\theta\int_0^{2\pi}\Phi\Phi^*d\varphi = 1. \qquad (28\cdot11)$$

Without loss of generality we can and will require that the separate factor integrals shall be normalized to unit values:

$$\int_0^\infty RR^*r^2dr = \int_0^\pi \Theta\Theta^* \sin\theta\, d\theta = \int_0^{2\pi}\Phi\Phi^*d\varphi = 1. \qquad (28\cdot12)$$

Adopting the exponential form of the Φ functions and introducing the appropriate value A, we obtain

$$\Phi_m(\varphi) = (2\pi)^{-\frac{1}{2}}\, e^{im\varphi}. \qquad (28\cdot13)$$

[1] COURANT-HILBERT, *M.M.P.*, pp. 272, 441; RIEMANN-WEBER, *D.P.*, Vol. I, p. 309.

28d. Determination of $\Theta(\theta)$ and Its Eigenvalues.—Having obtained
the eigenvalues of β, we can proceed to the solution of Eq. (28·8). Insert-
ing m^2 for β, and introducing x for $\cos\theta$ as a new independent variable,
we reduce the equation to the form

$$\frac{d}{dx}\left[(1-x^2)\frac{d\Theta}{dx}\right] + \frac{m^2}{x^2-1}\Theta + \lambda\Theta = 0. \tag{28·14}$$

This is to be solved in the interval $-1 < x < +1$ subject to the require-
ment (Sec. 17) that the product $R\Theta\Phi$ shall be continuous and twice
differentiable with respect to x, y, and z at every finite point except the
origin. We shall begin by seeking solutions which are merely continuous
everywhere (the line $\theta = 0,\pi$ contains the only points where continuity
comes in question) and will show later on that solutions of the differential
equation meeting this mild requirement are not only twice differentiable
but actually analytic everywhere except at the origin and at infinity.

In the special case that $m = 0$, Eq. (28·14) reduces to the equation of
the Legendre polynomials (27·1). The continuity requirement is
satisfied if we identify $\Theta(x)$ with one of these polynomials, say $P_l(x)$,
so that the functions $Y_{l,0}(\theta,\varphi) = P_l(\cos\theta)\Phi_0(\varphi)$ solve the problem of
Eq. (28·6) with the eigenvalues

$$\lambda = l(l+1). \qquad l = 0, 1, 2, 3, \cdots \tag{28·15}$$

We next search for solutions appropriate to other values of m. The
end points of the interval $-1 < x < +1$ are still regular singular points
and the fundamental indicial equation (26·5) for the point $x = +1$
reduces to $r^2 = (m/2)^2$. The roots of this equation differ by an integer
so that one solution of any fundamental system is sure to become infinite
at this point. The solutions which do not become infinite vanish at the
end points because the larger root of the indicial equation is positive.

A solution of Eq. (28·14) which satisfies the continuity requirement is
obtainable from each Legendre polynomial of order greater than, or
equal to, $|m|$. To prove this substitute $P_l(x)$ for y in equation (27·1)
and the corresponding eigenvalue $l(l+1)$ for λ. Differentiating with
respect to x and making the substitution

$$P_{l,1}(x) = (1-x^2)^{\frac{1}{2}} P_l'(x),$$

we obtain

$$\frac{d}{dx}\left[(x^2-1)\frac{dP_{l,1}}{dx}\right] - \frac{P_{l,1}}{x^2-1} - l(l+1)P_{l,1} = 0.$$

Repeating this process τ times we arrive at the equation

$$\frac{d}{dx}\left[(x^2-1)\frac{dP_{l,\tau}}{dx}\right] - \frac{\tau^2}{x^2-1}P_{l,\tau} - l(l+1)P_{l,\tau} = 0, \tag{28·16}$$

in which

$$P_{l,\tau}(x) \equiv (1-x^2)^{\frac{\tau}{2}}\frac{d^\tau}{dx^\tau}P_l(x) = \sin^\tau\theta\frac{d^\tau}{d(\cos\theta)^\tau}P_l(\cos\theta). \tag{28·17}$$

Since Eq. (28·16) is identical in form with Eq. (28·14) when τ is identified with $|m|$, the functions $P_{l,\tau}$ form eigenfunctions of our problem with the eigenvalues previously obtained for the case where m is zero, provided that they satisfy the boundary conditions.

As $P_{l,\tau}$ vanishes at $\theta = 0$ and $\theta = \pi$, it is evident at once that the product $Y_{l,m}(\theta,\varphi) = P_{l,|m|}(\cos \theta)\Phi_m(\varphi)$ is continuous over the entire sphere. Furthermore it is not difficult to prove that the product $r^l Y_{l,m}(\theta,\varphi)$ is a homogeneous polynomial of degree l in the Cartesian coordinates x, y, z.[1] As such it is an analytic function of these variables.

It follows that if $R(r)$ is any function of r analytic in the interval $0 < r < \infty$, the product $R(r)Y(\theta,\varphi)$ is an analytic function of x, y, z except at the origin and infinity. This means that the functions $Y_{l,m}(\theta,\varphi)$ satisfy all requirements.

The functions $P_{l,\tau}(x)$ are called *associated Legendre functions* and the products

$$Y_{l,m}(\theta,\varphi) = P_{l,\tau}(\cos \theta)\begin{Bmatrix}\sin m\varphi \\ \cos m\varphi\end{Bmatrix} \qquad \tau = |m|$$

are called *tesseral harmonics* of the lth degree and the mth order.[2] They form a special factorable type of surface harmonic.

28e. Completeness of System of Eigenfunctions.—Equation (28·14) is in the standard Sturm-Liouville form with $p(x)$ equal to $(1 - x^2)$ and $\rho(x)$ equal to unity. The solutions and their first derivatives are finite at the boundary points where $p(x)$ vanishes. Hence $P_{l,\tau}(x)$ satisfies the singular-point boundary conditions of Sec. 23d as well as the physical boundary conditions. Solutions of the differential equation which do not satisfy the s.p.b.c. are discontinuous at $x = \pm 1$. There is an interval G of λ values such that for each λ in the interval there is a pair of integral curves $u_\lambda(x)$ and $v_\lambda(x)$ which conform to the s.p.b.c. at $x = +1$ and $x = -1$, respectively. This interval extends from $-\infty$ to $+\infty$ and we infer from Sec. 25b that the totality of normalized eigenfunctions of Eq. (28·14) *for any fixed value of m*, solved subject to the s.p.b.c., forms a *complete orthogonal system*.

By Sec. 23 the nth eigenfunction in a series arranged according to increasing eigenvalues must have $n - 1$ nodes in the interior of the

---

[1] $r^{l-\tau}\dfrac{d^\tau}{d(\cos \theta)^\tau}P_l(\cos \theta)$ is a homogeneous polynomial of degree $l - \tau$ in x, y, z, being made up of terms of the form $(\cos \theta)^n r^n r^{l-\tau-n} = z^n(x^2 + y^2 + z^2)^{(l-\tau-n)/2}$. Furthermore

$$r^\tau(\sin \theta)^\tau e^{\pm i\tau\varphi} = (x \pm iy)^\tau$$

is a polynomial of degree τ. The product of these polynomials is a homogeneous polynomial of degree l, viz., $r^l Y_{l,m}(\theta,\varphi)$ if $m = \pm\tau$.

[2] *Cf.* W. E. BYERLY, *Fourier's Series and Spherical Harmonics*, pp. 195–199, Boston, 1893. Some authors define $P_{l,\tau}$ as $l!/(l + \tau)!$ times the value we have given.

fundamental interval $-1 < x < +1$. Hence, if we arrange the eigenfunctions we have already found in a similar series and discover that its nth member has no more than $n - 1$ nodes, we can be sure that we have found all the eigenfunctions there are. But it follows from the definition of $P_{l,\tau}(x)$ given in Eq. (28·17) that l cannot be less than τ. Hence the ordinal number of $P_{l,\tau}$ in such a series is $(l - \tau + 1)$, and since $d^\tau P_l(x)/dx^\tau$ is a polynomial of degree $l - \tau$, $P_{l,\tau}$ cannot have more than $l - \tau$ nodes in any interval. We conclude that *the series of eigenfunctions $P_{l,\tau}$ for $m = \pm\tau$ with $l = \tau, \tau + 1, \tau + 2, \cdots$ is complete.*

Degeneracy.—As the $(l + 1)$th derivative of $P_l(x)$ is zero, there are just $l + 1$ solutions of Eq. (28·8) corresponding to the same value of λ, viz., $\lambda_l = l(l + 1)$, and associated with the $l + 1$ permissible values of $\beta(= m^2)$. If β is zero, there is just one Φ function, viz., $\Phi_0 = (2\pi)^{-\frac{1}{2}}$. All other values of β yield two linearly independent Φ functions. Thus the lth eigenvalue of the Y equation, viz., $\lambda_l = l(l + 1)$, has $2l + 1$ linearly independent solutions which are factorable into products of the form $\Theta\Phi$. It is not difficult to show that the set of all possible product functions is complete, so that the total number of linearly independent solutions for λ_l is $2l + 1$. We say that λ_l has $(2l + 1)$-fold *degeneracy*.

The more important properties of the associated Legendre functions are summarized in Appendix F. Equation (F15) permits us to normalize the Θ functions. Denoting the normalized function derived from $P_{l,\tau}(x)$ by $\Theta_{l,m}$ we have[16]

$$\Theta_{l,m}(\theta) = \left[\frac{(2l + 1)(l - \tau)!}{2(l + \tau)!} \right]^{\frac{1}{2}} P_{l,\tau}(\cos\theta), \qquad \tau = |m|. \quad (28\cdot18)$$

28f. Physical Interpretation of Quantum Numbers l and m.—We have now located the eigenvalues of λ and determined the associated angle functions. We proceed to the study of the radial wave equation (28·5). Substituting the value of λ from Eq. (28·15) and changing the dependent variable from R to $\Re = rR$ [cf. Eqs. (23·19) and (23·20)], we obtain

$$\frac{d^2\Re}{dr^2} + \kappa\left[E - V(r) - \frac{l(l + 1)}{\kappa r^2} \right]\Re = 0, \quad (28\cdot19)$$

an equation identical in form with (18·2), but with an effective potential energy

$$V_e(r) = V(r) + \frac{l(l + 1)}{\kappa r^2} \quad (28\cdot20)$$

substituted for $V(r)$. If we make the further substitution

$$\Re = e^{\frac{2\pi i S_r}{h}}$$

[1] Graphs of the normalized Θ's are given in Condon and Morse, *Quantum Mechanics*, Fig. 6, p. 55.

and the approximations of Sec. 11, we derive the relation

$$\frac{1}{2\mu}\left(\frac{dS_r}{dr}\right)^2 + V(r) + \frac{l(l+1)}{\kappa r^2} = E, \tag{28·21}$$

which is the same as the classical *radial component* of the Hamilton-Jacobi equation with $l(l+1)h^2/(4\pi^2)$ substituted for the square of the classical angular momentum. This result suggests that in the wave mechanics $\frac{h}{2\pi}\sqrt{l(l+1)}$ is the value of the angular momentum, an interpretation which will be justified in the next chapter from another point of view. As $l(l+1) = (l+\frac{1}{2})^2 - \frac{1}{4}$, this expression is very similar to the expression $(l+\frac{1}{2})\frac{h}{2\pi}$ for the angular momentum in the Bohr theory with half-integral quantum numbers, but permits a state of zero angular momentum forbidden in the latter theory. In view of this partial agreement with the Bohr theory we adopt the terminology of that theory and refer to the integer l as the *azimuthal quantum number*. (In the Bohr theory l is associated with the azimuthal angle in the plane of the classical orbit.)

A similar relationship exists between the integer m of the present theory and the so-called "magnetic quantum number" of the Bohr theory. The latter, when multiplied by $h/(2\pi)$, represented the component of angular momentum in the direction of the z axis which was taken to be that of an incipient magnetic field. Hence the ratio $\frac{m}{l+\frac{1}{2}}$ or m/k gave the cosine of the angle between the normal to the orbit and the z axis. The angle θ between the radius vector and the z axis for such a Bohr orbit ranges accordingly from $-\sin^{-1}\left(\frac{m}{l+\frac{1}{2}}\right)$ to $+\sin^{-1}\left(\frac{m}{l+\frac{1}{2}}\right)$. The same sharply defined range of values for θ is obtained from the present theory, if one uses large values of l. If $l-m$ is large, the nodes of Θ are closely spaced and it is convenient to adopt the procedure of Sec. 21i, replacing the probability $\Theta^2 \sin\theta\, d\theta$ that the particle lies in the range $d\theta$ by its mean value over the interval between successive nodes. In the limit the probability is[1]

$$\overline{\Theta^2} \sin\theta\, d\theta \cong \frac{\text{constant} \times |\sin\theta|d\theta}{\sqrt{\sin^2\theta - \frac{(m^2 - \frac{1}{4})}{(l+\frac{1}{2})^2}}}, \qquad \text{if } \sin^2\theta > \frac{m^2 - \frac{1}{4}}{(l+\frac{1}{2})^2}$$

$$\cong 0, \qquad\qquad \text{if } \sin^2\theta < \frac{m^2 - \frac{1}{4}}{(l+\frac{1}{2})^2}.$$

[1] Changing the dependent variable in (28·8) from Θ to U, where $U = \sqrt{\sin\theta}\,\Theta$, we obtain the differential equation

Figure 9 shows $\Theta^2 \sin \theta$ plotted against θ for the case $l = 10$, $m = 5$, and for comparison the smoothed function $\overline{\Theta^2} \sin \theta$. The concentration of probability in a limited range of polar angles is, of course, greatest when $m = l$ so that Θ^2 is proportional to $(\sin \theta)^{2l}$.
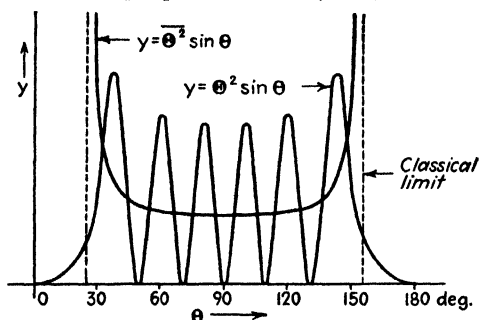
Fig. 9.—Probability distribution for the polar angle θ.

28g. Behavior of Radial Wave Functions at Boundary Points.— Before introducing an explicit potential function let us examine the behavior of the integral curves of the radial wave equation (28·19) at the boundary points and verify the equivalence of the physical boundary conditions and the s.p.b.c. for this Sturm-Liouville problem. We assume that $V(r)$ approaches a definite limit V_∞ as r becomes infinite, and that for small values of r the Coulomb inverse square law of force applies. Then

$$V = \frac{\alpha}{r} + \varphi(r), \tag{28·22}$$

where α is a constant and $\varphi(r)$ is regular at $r = 0$.

$$\frac{d^2 U}{d\theta^2} - \left(\frac{\beta - \frac{1}{4}}{\sin^2 \theta}\right) U + \left(\lambda + \frac{1}{4}\right) U = 0.$$

This is of the form of (18·2) so that an approximate solution for large values of λ is obtainable by the B. W. K. method. The "wave length" of U with respect to the independent variable θ is $2\pi|\sin \theta| \div [(\lambda + \frac{1}{4}) \sin^2 \theta - \beta + \frac{1}{4}]^{\frac{1}{2}}$. We denote this quantity by $\mu(\theta)$. It becomes imaginary when $\sin^2 \theta < \dfrac{\beta - \frac{1}{4}}{\lambda + \frac{1}{4}}$. Hence U is oscillatory when $\sin^2 \theta > \dfrac{\beta - \frac{1}{4}}{\lambda + \frac{1}{4}} = \dfrac{m^2 - \frac{1}{4}}{(l + \frac{1}{2})^2}$, and fades abruptly to zero outside this limit. The B. W. K. approximation (21·7) takes the form

$$U \cong \text{constant} \times \sqrt{\mu(\theta)} \, \cos\left(2\pi \int \frac{d\theta}{\mu(\theta)}\right).$$

Replacing $\left(\cos 2\pi \int \dfrac{d\theta}{\mu}\right)^2$ by its average value for the interval between successive nodes, one obtains

$$\overline{\Theta^2} \sin \theta \cong \overline{U^2} \cong \frac{\text{constant} \times \sin \theta}{\sqrt{\sin^2 \theta - \dfrac{(m^2 - \frac{1}{4})}{(l + \frac{1}{2})^2}}}.$$

If E is greater than V_∞ the function $\mathfrak{R} = Rr$ of Eq. (28·19) will approximate a sine curve for large values of r and cannot be made quadratically integrable (*cf.* Sec. 19c, p. 85). Neither the physical boundary conditions nor the s.p.b.c. can be satisfied and we conclude that there are no type A eigenfunctions for energies greater than V_∞.

Next assume that $E < V_\infty$. It follows from Sec. 19 and Appendix C that for each such value of E there are integral curves of Eq. (28·19) which, together with their derivatives, approach zero monotonically and "exponentially" as r becomes infinite. Such integral curves satisfy the s.p.b.c. at infinity and yield solutions of Eq. (28·5) by the transformation $R = \mathfrak{R}/r$ which again satisfy the s.p.b.c. at infinity. As the transforms $R(r)$ are bounded in the neighborhood of infinity they also satisfy the physical boundary conditions for type A functions. Conversely, any integral curve which does *not* conform to the type A conditions at infinity will not be quadratically integrable and *cannot* satisfy the s.p.b.c. at the outer boundary. We conclude that in the neighborhood of $r = \infty$ the s.p.b.c. applied to solutions of either of the radial equations (28·5) and (28·19) are equivalent to the boundary conditions for type A functions.

Let us next apply the theory of Sec. 26 to the inner boundary point. The point $r = 0$ is again a regular singular point. Using the symbol t for the exponent of the lowest power of r in $\mathfrak{R}(r)$, the indicial equation [*cf.* Eq. (26·5)] of (28·19) at the origin is

$$t(t - 1) - l(l + 1) = 0.$$

Its roots are $t_1 = l + 1$ and $t_2 = -l$. As these two roots differ by an integer we have to do with the logarithmic case of Eq. (26·6). The two linearly independent solutions given by Eqs. (26·3) and (26·6) yield

$$R_1 = \frac{\mathfrak{R}_1}{r} = r^l f_1(r),$$

$$R_2 = \frac{\mathfrak{R}_2}{r} = A r^l f_1(r) \log r + r^{-(l+1)} f_2(r),$$

where f_1 and f_2 are regular at $r = 0$ and do not vanish there.

R_1 and $\mathfrak{R}_1$ satisfy the singular-point boundary conditions for the equations (28·5) and (28·19), respectively—also the physical boundary conditions—while R_2 satisfies neither the s.p.b.c. nor the physical conditions. We conclude that in this case the physical boundary conditions for type A functions are equivalent to the s.p.b.c. at both inner and outer boundary points. The range G of energy values over which these conditions can be met at both ends of the fundamental interval extends from $E = -\infty$ to $E = V_\infty$. The attention of the reader is called to the fact that *both* R_2 and R_1 are quadratically integrable when $l = 0$. Hence,

the mere requirement of quadratic integrability would *not* suffice to pick out the desired orthogonal set of wave functions.

Let n_l'' denote the number of interior nodes of an integral curve of (28·19) for the energy $E = V_\infty$ and for the azimuthal quantum number l. Then it follows from Sec. 23 that there will be n_l'' discrete eigenvalues of E below V_∞. n_l'' may be zero in the case of an effective potential energy with a very shallow minimum or none at all, or it may be infinite. The Sommerfeld phase integral of Eq. (21·10) evaluated for $E = V_\infty$ gives an approximate means for evaluating n_l'' from which we see that if $V - V_\infty$ is negative and varies as $1/r$ for large values of r, the number of discrete eigenvalues is infinite independent of the azimuthal quantum number l. On the other hand n_l'' is finite and varies with l if $V - V_\infty$ varies as $1/r^2$ or some higher inverse power of r in the neighborhood of infinity.

Regardless of the number of eigenvalues, the sequence has an upper bound V_∞, and hence *the eigenfunctions do not form a complete system.*

In applications of the general theory to the vibration and rotation of diatomic molecules and to single-electron atoms the number of interior nodes of the radial eigenfunction is called the *radial* or *vibrational quantum number*. In molecular theory it is usually designated by the symbol v which we here adopt. $v + 1$ is the ordinal number of any eigenvalue of the energy when arranged in an ascending series together with the other eigenvalues for the same azimuthal quantum number l. In one-electron atomic problems the so-called "total quantum number"

$$n \equiv v + l + 1$$

is introduced instead of v because of its major influence on the energy. The associated Legendre functions Θ_{lm} have $l - m$ nodes and the real and imaginary parts of Φ_m have each $2m$ nodes. However, if the spherical harmonic $Y_{l,m} = \Theta_{lm}\Phi_m$ is laid out on the sphere we see that the $2m$ nodes in the real part of Φ_m may be united in pairs to form m nodal meridian planes. Thus $Y_{l,m}$ may be said to have l nodes in all, while the complete wave function

$$\psi_{nlm} = R_{nl}\Theta_{lm}\Phi_m$$

has $v + l = n - 1$ nodes. These divide the space around the origin into cells in each of which the phase of the vibration for, say, the real part of ψ, is opposite to that of its neighbors.

As previously noted there are $2l + 1$ linearly independent factorable solutions of the Y equation (28·6) for each value of l. This degeneracy carries over to the general equation (28·3) which has $2l + 1$ linearly independent factorable solutions for each energy level. It is due to the spherical symmetry of the problem and corresponds exactly to the multiplicity of orbital orientations and azimuthal quantum numbers in

the Bohr theory for the same system. As the orientation of the axes in space is arbitrary, the independent factorable solutions may be chosen in an infinity of different ways.

As explained in Sec. 21*h*, we can compute approximate eigenvalues of Eq. (28·19) by the B. W. K. method, provided that we replace the potential function $V(r)$ by $V(r) + \dfrac{h^2}{32\pi^2\mu r^2}$. The quantum condition (21·10) then takes the form

$$J(E,l) \equiv \oint \sqrt{2\mu\left[E - V(r) - \frac{(l + \frac{1}{2})^2}{\kappa r^2}\right]}\,dr = (n + \frac{1}{2})h.$$

$$n = 0,1,2, \cdots \quad (28\text{·}23)$$

In other words we compute the radial momentum as if the angular momentum were given the value $(l + \frac{1}{2})h/2\pi$ instead of $[l(l + 1)]^{\frac{1}{2}}h/2\pi$. Thus it is to be expected on the basis of the work done in Sec. 21 that the Bohr energy-level formulas will give the correct energies to a close approximation provided that we use half-integral values of the azimuthal and radial, or vibrational, quantum numbers.

There are two important special cases to be considered, *viz.*, (*a*) the problem of the hydrogenic, or one-electron, atoms—to be discussed in Sec. 29—and (*b*) the problem of the nongyroscopic diatomic molecule. The latter system is obviously favorable to the application of the B. W. K. method on account of the large reduced mass of every diatomic molecule as compared with the electronic mass to be used in atomic problems. (The accuracy of the usual first B. W. K. approximation is inversely proportional to the integral μr of Eq. (21·20) which, in turn, varies inversely with the square root of the mass.)

28h. The Dumbbell Model of the Diatomic Molecule.—In applying the theory of the two-body problem to a diatomic molecule we make use of what is called the "dumbbell" mathematical model. It neglects the details of the electronic structure entirely, assuming that the only effect of the electrons on the nuclear motion is through a modification which they produce in the effective mutual potential energy of the two heavy nuclei. The use of this model is justifiable on theoretical grounds for those "electronic states" of the molecule which have no average orbital electronic angular momentum about the internuclear axis (*cf.* Sec. 47*c*).

Of course it is not possible to work out a definite formula for molecular energy levels from Eq. (28·23) without definite information regarding the potential function $V(r)$, and just this essential information is lacking. In principle, $V(r)$ is theoretically predictable on the basis of a solution of the energy-level problem for the electrons moving under the influence of fixed nuclei. In practice we are unable to compute $V(r)$ in this way with

any degree of accuracy except in the cases of neutral and ionized molecular hydrogen. Hence $V(r)$ must be regarded as *a priori* a quantitatively unknown function. From a qualitative point of view, however, we can safely predict that $V(r)$ will have the general form shown in Fig. 10 with a single minimum at the internuclear distance r_0. The character of the curve shown in the figure is dictated by the following considerations.

a. At very short distances the electronic contribution to $V(r)$ must approach a finite limit which can be identified with the electronic energy of a single atom formed by uniting the two nuclei. As the contribution to $V(r)$ from the Coulomb repulsion of the bare nuclei becomes infinite

Fig. 10.—Potential energy and energy eigenvalues for model of diatomic molecule.

at $r = 0$, $V(r)$ itself must have a pole of the first order at the origin. Thus $\lim_{r \to 0} rV(r) = Z_1 Z_2 e^2$, where Z_1 and Z_2 are the atomic numbers of the two nuclei.

b. At very large distances the force between two neutral atoms must approach zero more rapidly than r^{-2}, and $V(r)$ itself must approach a finite limit which can be set equal to zero without loss of generality. If a stable molecule is to be possible $V(r)$ must have a minimum at some intermediate point $r = r_0$ of the order of magnitude of an atomic diameter. The absolute value of $V(r_0)$ constitutes what we may describe as the static dissociation energy of the molecule and is very nearly equal to the dissociation energy measured by chemical methods.

c. It seems reasonable to assume that $V(r)$ will be a relatively smooth curve with only the one minimum. This postulate finds support in approximate computations and in the analysis of empirical band spectra.

The application of the quantum condition (28·23) to the evaluation of any energy level requires an accurate specification of the potential function $V(r)$ only over the classical range of vibration for the corresponding energy. Hence it suffices for the determination of the lower energy levels of the molecule that we have a description of $V(r)$ which is accurate in the neighborhood of the minimum point r_0. A power series in $r - r_0$, or $\xi \equiv \dfrac{r - r_0}{r_0}$, is obviously adapted to our purpose and in favorable cases should converge rapidly over the required range of values. It is possible to insert such a series, say

$$V(r) = -V(r_0) + k\xi^2(1 + a\xi + b\xi^2 + \cdots),$$

into (28·23) and to evaluate $J(E,l)$ as a power series in $E - V(r_0)$. Reversing the latter series, one obtains an expression for E as a double

power series in J and l, the coefficients being known functions of $r_0, a, b,$
$\cdots$.[1] Inserting $(v + \frac{1}{2})h$ for J, and dropping higher power terms, we
obtain in crude first approximation

$$E = V(r_0) + (v + \tfrac{1}{2})h\omega_0 + (l + \tfrac{1}{2})^2 \frac{h^2}{8\pi^2\mu r_0{}^2} + \cdots$$

$$v = 0, 1, 2, \cdots, \qquad l = 0, 1, 2, \cdots \qquad (28\cdot24)$$

where ω_0 is the classical vibrational frequency of the molecule for small

oscillations about r_0, viz., $\dfrac{1}{2\pi}\sqrt{\dfrac{k}{\mu}}$. The corresponding approximation for

the integral form of Bohr theory is

$$E = V(r_0) + vh\omega_0 + \frac{l^2h^2}{8\pi^2\mu r_0{}^2} + \cdots \qquad \begin{cases} v = 0, 1, 2, \cdots \\ l = 1, 2, 3, \cdots \end{cases} \quad (28\cdot25)$$

What may be described as the rotational spacing of the energy levels
is obtained by holding v fixed and allowing l to take on successive integral
values. According to (28·24) the successive intervals obtained in this
way are in the ratio 2, 4, 6, $\cdots$, whereas (28·25) gives intervals in the
ratio 3, 5, 7, $\cdots$. Observations on band spectra show that the former
ratios are correct, at least for small values of l.

The vibrational spacing obtained by holding l fast and varying v is
the same for both formulas and gives us no hold on the legitimacy of half-
integral vibrational quantum numbers. However, if we take the energy
difference between two levels belonging to different electronic states
but with v and l in each case given their lowest values, we obtain a con-
tribution $h\omega_0/2$ from (28·24) which is absent when we use (28·25). This
contribution shows up when we compare the spectra of two molecules
with chemically similar atoms but different nuclear masses. Thus it
enters into the isotope effect in band spectra which has been shown by
Mulliken to require the use of half-integral vibrational quantum numbers.

To summarize, we may say that in the study of the rotational and
vibrational levels of diatomic molecules we have a field in which wave
mechanics gives essentially the same predictions as a modified Bohr
theory with half-integral quantum numbers. These predictions are
verified experimentally.

29. THE HYDROGENIC ATOM

29a. Application of the B. W. K. Method.—The most important two-
body problem is that of the hydrogenic atoms. We assume a nucleus of

---

[1] *Cf.* E. C. KEMBLE, *Proc. Nat. Acad. Sci.* **7**, 283 (1921); *J.O.S.A.*, **12**, 1 (1926)
Another and equivalent procedure due to Kratzer is given in A. E. Ruark and H. C.
Urey, *Atoms, Molecules, and Quanta*, sec. 4, Chap. XII, New York, 1930, and elsewhere.
Cf. also J. L. Dunham, *Phys. Rev.* **41**, 721 (1932) for higher B. W. K. approximations.
The author regards the criticism of the B. W. K. method by Rosenthal and Motz,
Proc. Nat. Acad. Sci. **23**, 259 (1937) as unwarranted.

charge Ze and a single electron under the influence of a Coulomb potential $-e^2Z/r$.

The B. W. K. method yields particularly good results in this case in spite of the small electronic mass, because the effective potential energy for the radial motion consists of two terms, one varying as r^{-1} and the other as r^{-2}. It follows that the classical local momentum p has no zeros except those at the classical turning points. Hence we can choose the path Γ of Sec. 21h as the sum of a portion of the axis of imaginaries and a quarter circle of arbitrarily large radius drawn about the origin as a center. In this case μ_Γ reduces to the very small contribution involved in getting away from the origin and the *a priori* accuracy of the procedure is high. Actually, the B. W. K. energy values are identical with those derived from an exact integration of the radial equation.

The phase integral $J(E,l)$ of Eq. (28·23) can be evaluated explicitly.[1] We obtain

$$J(E,l) = -(l + \tfrac{1}{2})h - \frac{2\pi\mu Ze^2}{\sqrt{-2\mu E}} = (v + \tfrac{1}{2})h.$$

Solving for E, one derives the familiar formula

$$E = -\frac{2\pi^2\mu Z^2 e^4}{n^2 h^2} = -\frac{NZ^2 hc}{n^2}, \tag{29·1}$$

where N is the Rydberg constant in wave-number units and n is an abbreviation for the integer $v + l + 1$. The combination of the half-integral radial quantum number $v + \tfrac{1}{2}$, with the azimuthal quantum number l, and the correction of the B. W. K. method for radial motion, to give the simple formula (29·1), with its integral denominator, is rather curious.

The same result is readily obtainable along with the corresponding eigenfunctions by exact integration of the differential equation for $\mathcal{R}(r)$. This equation was rigorously solved in Schrödinger's first paper and gave the initial impetus to this point of view. He treated the equation by a method of complex integration due to Poincaré and Horn. The more elementary Sommerfeld polynomial method of getting the solution is described herewith.

29b. Application of Polynomial Method.—As a first step we express the energy E and the radius r in terms of new units. As unit energy we choose the value $2\pi^2\mu e^4 Z^2/h^2$ of the lowest energy level in the Bohr theory. As unit radius we choose the radius of the innermost Bohr orbit, *viz.*, $h^2/4\pi^2\mu e^2 Z = a_0/Z$. Then

[1] *Cf.* A. Sommerfeld, *Ann. d. Physik* **51**, 1 (1916), any edition of Sommerfeld's *Atombau und Spektrallinien,* or Ruark and Urey, *Atoms, Molecules, and Quanta,* Chap. V, New York, 1930.

$$E \rightarrow \epsilon = \frac{Eh^2}{2\pi^2 \mu e^4 Z^2} = \frac{a_0^2 \kappa E}{Z^2};$$

$$r \rightarrow \rho = \frac{4\pi^2 \mu e^2 Z}{h^2} r = \frac{rZ}{a_0}.$$

The differential equation takes the simplified form

$$\frac{d^2 \mathfrak{R}}{d\rho^2} + \left[\epsilon + \frac{2}{\rho} - \frac{l(l+1)}{\rho^2} \right] \mathfrak{R} = 0. \tag{29.2}$$

Following the procedure of Sec. 20 we drop the terms in ρ^{-1} and ρ^{-2} in order to examine the behavior of the solutions for very large values of ρ. We find that

$$\mathfrak{R}(\rho) \cong e^{\pm \sqrt{-\epsilon}\,\rho}. \qquad \rho \gg 1$$

In order to satisfy the boundary conditions it is necessary that ϵ shall be negative. For the same reason the solution with the positive exponent must be rejected. To obtain a solution valid for small r we make the substitution

$$\mathfrak{R} = \rho^{l+1} e^{-\sqrt{-\epsilon}\,\rho} L(\rho), \qquad \rho \ll 1$$

and obtain the differential equation

$$\rho L'' + 2[l + 1 - \rho\sqrt{-\epsilon}]L' + 2[1 - (l+1)\sqrt{-\epsilon}]L = 0 \tag{29.3}$$

for L. L should be regular at the origin, and we therefore assume the power series

$$L = \sum_{\nu=0}^{\infty} a_\nu \rho^\nu, \tag{29.4}$$

which yields the recurrence formula

$$a_{\nu+1}[(\nu+1)(2l+\nu+2)] = 2a_\nu[(l+\nu+1)\sqrt{-\epsilon} - 1]. \tag{29.5}$$

The series breaks off with the νth term and yields a solution of the boundary, value problem if

$$\epsilon = -\frac{1}{(\nu+l+1)^2} = -\frac{1}{n^2}. \tag{29.6}$$

The corresponding energy values are given by (29.1). If ϵ does not meet the condition (29.5), the series for $L(\rho)$ behaves like $\rho^{-(l+1)}e^{2\sqrt{-\epsilon}\,\rho}$ for large values of ρ, while that for $\mathfrak{R}$ behaves like $e^{\sqrt{-\epsilon}\,\rho}$. Hence $\mathfrak{R}$ becomes infinite at $\rho = \infty$. We conclude that the solutions obtained above are the only solutions which satisfy the boundary conditions. We may draw the same inference from the fact that the factor $L_{\nu l}(\rho)$, in the νth eigenfunction for any given l, is a polynomial of degree ν and hence cannot have more than ν nodes (cf. Sec. 28e).

29c. Generalized Laguerre Polynomials.—The polynomials $L_{vl}(\rho)$ defined by Eqs. (29·3) and (29·4) are identical to a constant factor with the generalized Laguerre polynomials $L_{n+l}{}^{2l+1}(2\rho/n)$ described in Appendix G. This may be proved by using the substitution

$$x = \frac{2\rho}{n} = \frac{2r}{na_0}Z$$

to reduce Eq. (29·2) to the form of Eq. (G-8), Appendix G. The radial eigenfunctions now take the form

$$\mathcal{R}_{nl}(r) = C_{nl}x^{l+1}e^{-\frac{x}{2}}L_{n+l}{}^{2l+1}(x). \tag{29·7}$$

The constant C_{nl} can be determined to satisfy the normalization condition (28·12). Then Eq. (G-12) of Appendix G yields

$$\int_0^\infty RR^*r^2dr = \frac{na_0}{2Z}\int_0^\infty \mathcal{R}_{nl}{}^2dx = \frac{na_0}{2Z}C_{nl}{}^2\frac{2n[(n+l)!]^3}{(n-l-1)!} = 1,$$

and the normalized radial wave function is

$$R_{nl}(r) = \sqrt{\frac{(n-l-1)!}{2n[(n+l)!]^3}\left(\frac{2Z}{na_0}\right)^3}\left(\frac{2rZ}{na_0}\right)^l e^{-\frac{rZ}{na_0}}L_{n+l}{}^{2l+1}\left(\frac{2rZ}{na_0}\right) \tag{29·8}$$

Multiplying together the different factors of our complete normalized wave functions, we obtain

$$\psi_{nlm}(r,\theta,\varphi) = \left[\left(\frac{2Z}{na_0}\right)^3\left(\frac{2l+1}{4\pi}\right)\frac{(l-|m|)!(n-l-1)!}{(l+|m|)!2n[(n+l)!]^3}\right]^{1/2}e^{im\varphi}$$

$$\times P_{l,|m|}(\cos\theta)\left(\frac{2rZ}{na_0}\right)^l e^{-\frac{rZ}{na_0}}L_{n+l}{}^{2l+1}\left(\frac{2rZ}{na_0}\right). \tag{29·9}$$

Tables containing the explicit forms of the radial wave functions for the different lower energy levels are given in Pauling and Goudsmit, *The Structure of Line Spectra*, New York, 1930. In the case of the lowest energy level (normal state) of an hydrogenic atom we have the single wave function $\psi_{100} = \pi^{-1/2}\left(\dfrac{Z}{a_0}\right)^{3/2}e^{-\frac{rZ}{a_0}}$.

As the energy is fixed by the single quantum number $n(= v + l + 1)$, the hydrogen atom in the present approximation (neglecting relativity and spin corrections) has a special degeneracy referable to the Coulomb law of force which does not occur in the general two-body problem. The energy level E_n has l values ranging from 0 to $n-1$. Hence the number of different sets of values of l and m, i.e., the number of independent factorable solutions, is

$$\sum_{l=0}^{n-1}(2l+1) = n^2.$$

29d. The Most General Eigenfunction.—The most general possible type A wave function for the energy level E_n is a linear combination of the n^2 terms of the type $R_{nl}\Theta_{lm}\Phi_m$. To prove the point we assume that $\psi_n(r,\theta,\varphi)$ is a type A eigenfunction corresponding to the energy E_n. ψ must then be periodic in φ with the period 2π. Hence we may expand it into the complex Fourier's series

$$\psi_n = \sum_{-\infty}^{+\infty} F_m(r,\theta)e^{im\varphi}.$$

As the functions Θ_{lm}, for any given value of m, form a complete orthogonal series, we can expand each of the F's in terms of them. Then

$$\psi_n = \sum_{m=-\infty}^{+\infty} \sum_{l=|m|}^{\infty} G_{lm}(r)\Theta_{lm}(\theta)e^{im\varphi}.$$

If we substitute this series into Eq. (28·3) and make use of the fact that Θ_{lm} and $e^{im\varphi}$ are solutions of Eqs. (28·8) and (28·9), respectively, we obtain

$$\sum_{m=-\infty}^{+\infty} \sum_{l=|m|}^{+\infty} \Theta_{lm}e^{im\varphi}\left\{\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dG_{lm}}{dr}\right) + \kappa\left[E_n - V - \frac{l(l+1)}{\kappa r^2}\right]G_{lm}\right\} = 0.$$

Let this equation be multiplied through by $\Theta_{l'm'}e^{-im'\varphi}$ and integrated term by term over all values of θ and φ. Most of the terms drop out on account of the orthogonality properties of the Θ and Φ functions and the equation reduces to

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dG_{l'm'}}{dr}\right) + \kappa\left[E_n - V - \frac{l'(l'+1)}{\kappa r^2}\right]G_{l'm'} = 0. \qquad (29\cdot10)$$

Equation (29·10) has the form of Eq. (28·5) with $\lambda = l'(l'+1)$. It reduces to the form (29·2) if we insert the expression $-Ze^2/r$ for V, introduce appropriate units, and change the dependent variable from G to $\mathfrak{R} = rG$. Hence it has no nontrivial quadratically integrable solutions unless $n = l' + v + 1$, where v is an integer greater than or equal to zero. This means that all the G's vanish for which $l' > n - l$. The G's which do not vanish identically are the radial eigenfunctions $R_{nl'}$ previously derived. Thus the theorem is proved.

The space factor obtained in this manner is conveniently written in the form

$$\psi_n = \sum_{l=0}^{n-1} \sum_{m=-l}^{+l} c_{nlm}R_{nl}\Theta_{lm}\Phi_m. \qquad (29\cdot11)$$

If ψ_n, R_{nl}, Θ_{lm}, Φ_m are normalized,

$$\sum_{l=0}^{n-1} \sum_{m=-l}^{+l} |c_{nlm}|^2 = 1.$$

CHAPTER VI

THE CONTINUOUS SPECTRUM AND THE BASIC PROPERTIES OF SOLUTIONS OF THE MANY-PARTICLE PROBLEM

30. THE CONTINUOUS SPECTRUM IN ONE-DIMENSIONAL PROBLEMS

30a. The Nature and Use of the Eigenfunctions of the Continuous Spectrum.—We have already seen in Sec. 19 that if the potential-energy function $V(x)$ in a linear oscillator problem approaches a finite limit D as the coordinate x becomes infinite, the eigenvalue problem has no type A solutions for energies greater than D. For every value of E greater than D, however, there is another type of eigenfunction (type B) which is bounded and continuous everywhere, but not quadratically integrable. Classical energies greater than D correspond to *aperiodic motions* in which the particle does not remain permanently in the neighborhood of the potential minimum but flies off to infinity after its first collision with the potential barrier on the left (*cf.* Fig. 6). As all energies greater than D are on the same footing, the problem of finding the allowed values of E in this range does not have to be solved. We shall nevertheless be interested in wave packets whose average energy is greater than D. In order to develop a proper theory of the relative probabilities of different energies for such packets we should like to express arbitrary physically·admissible solutions of $H\Psi = -\dfrac{h}{2\pi i}\dfrac{\partial \Psi}{\partial t}$ as linear combinations of single-energy functions. Resolutions of this type are also of considerable importance for perturbation theory.

The desired resolution cannot be obtained with the aid of any linear combination of distinct type B eigenfunctions, for such a combination is never quadratically integrable. The difficulty can be avoided, however, by using an integral over a continuous family of the continuous-spectrum eigenfunctions. We are already familiar with an example of the appropriate type of analysis through our application of the Fourier integral theorem to the wave packets of free particles. The integrands of Eqs. (9·4) and (10·3) are actually type B eigenfunctions of $H\psi = E\psi$ for free particles in one dimension and three dimensions, respectively.

The theory of the continuous spectrum is nevertheless appreciably more difficult to handle than the theory of discrete spectra. Hence it is of importance to note that for many purposes some or all of the difficulties can be sidestepped by making use of the fact that by a slight modification of many physical problems, of no practical importance as regards

162

the final outcome, one can substitute a purely discrete spectrum for the discrete-continuous one. For this purpose one has only to place the system under discussion in a large imaginary box, requiring that Ψ shall vanish whenever one of the particles which compose the system touches its surface. It is evident that if the box is large enough, it will be legitimate to assume that results computed for the modified problem are as good as those computed for the original one in which coordinate space extends to infinity. The box eliminates the continuous spectrum completely both in one-dimensional and many-dimensional problems.

30b. The Weyl Theory.—In Chap. IV we have developed the required theory of the discrete eigenfunctions of the Sturm-Liouville equation

$$\Lambda y = \frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) - q(x)y = -\lambda\rho(x)y \tag{30.1}$$

for the case in which the fundamental interval is bounded by singular points a and b. In Sec. 23 we determined the properties of solutions of (30.1) in the neighborhood of the boundary points in order that there shall be a discrete spectrum of eigenvalues extending through a finite, or infinite, interval of real values of λ. In Sec. 25 we proved the completeness of the system of discrete eigenfunctions for the case in which the discrete spectrum extends to $+\infty$ and so laid the foundation for the expansion theorem. What we now desire is a generalization of these results for cases in which the discrete spectrum does not extend to infinity.

The problem has been treated by Weyl[1] in a basic paper which unfortunately involves an elaborate mathematical technique and makes difficult reading for the non-specialist. The class of problem discussed explicitly by Weyl is that in which (30.1) is to be solved in the interval $0 \leqslant x < \infty$, the left-hand boundary point being nonsingular and the boundary conditions applied at this point being of the homogeneous type. It is clear, however, that the validity of his conclusions will not be affected if the left-hand boundary is at an arbitrary singular point a, provided that we use the singular-point boundary condition at a and provided also that for any real λ a unique choice of the non-multiplicative constant of integration gives a solution of (30.1) which conforms to those boundary conditions and has a finite number of nodes in the neighborhood of a. We assume the same restrictions on the coefficients $p(x)$, $q(x)$, $\rho(x)$ as before.[2]

[1] Hermann Weyl, *Math. Annalen* **68**, 220 (1910). Important contributions to the theory of the continuous spectrum have also been made by E. Fues, *Ann. d. Physik* (4) **81**, 281 (1926) and J. R. Oppenheimer, *Zeits. f. Physik* **41**, 268 (1927), *Phys. Rev.* **31**, 66 (1928).

[2] The right-hand boundary point can also be brought in from infinity. but for our purpose this is unnecessary.

Let us next assume that the real values of λ fall in one of two intervals Ω_1 and Ω_2 having the following properties. In Ω_1 the right-hand boundary at $x = \infty$ has the same properties as the left-hand boundary. This is then a region in which a discrete spectrum can exist. In Ω_2 it is not possible to satisfy the singular-point boundary conditions at $x = +\infty$ at all.

Under the circumstances described we can deduce from Weyl's paper and Sec. 23 that in the interval Ω_1 there will be a normal discrete spectrum or no spectrum of eigenvalues at all. In Ω_2 there is a continuous spectrum of eigenvalues whose (type B) eigenfunctions satisfy the s.p.b.c. at $x = a$ and yield so-called *eigendifferentials* which satisfy them both at $x = a$ and at $x = \infty$.

In order to define the eigendifferentials we assume that $y(x,\lambda)$ denotes a real continuous function of x and λ which for every fixed λ in Ω_2 is a solution of (30·1) conforming to the s.p.b.c. at $x = a$. It includes all suitably normalized type B eigenfunctions of the equation. Then the function

$$\Delta_r y = \int_{\lambda_r}^{\lambda_r + \eta_r} y(x,\lambda)d\lambda \qquad (30\cdot 2)$$

is called an *eigendifferential* if the interval $\lambda_r < \lambda < \lambda_r + \eta_r$ lies in Ω_2. The fact that $\Delta_r y$, unlike $y(x,\lambda)$, satisfies the s.p.b.c. at $x = \infty$ for arbitrarily small values of η_r is the criterion by which we identify $y(x,\lambda_r)$ as a type B eigenfunction and λ_r as a continuous-spectrum eigenvalue.

Let us now assume that $f(x)$ is an arbitrary twice-differentiable function of x which conforms to the s.p.b.c. at $x = a$ and yields convergent integrals

$$\int_a^\infty \rho(x)|f|^2 dx; \qquad \int_a^\infty |\Lambda f|^2 \rho^{-1}dx. \qquad (30\cdot 3)$$

Then $f(x)$ has an absolutely and uniformly convergent expansion of the form

$$f(x) = \sum_{\text{all } n} c_n y_n(x) + \int_{\Omega_2} c(\lambda)y(x,\lambda)d\lambda. \qquad (30\cdot 4)$$

Here the discrete eigenfunctions $y_n(x)$ are conveniently normalized in the usual way, but the type B eigenfunctions require a different normalization usually stated as the condition that[1]

$$\lim_{\eta \to 0}\left[\eta^{-1}\int_0^\infty dx\rho(x)\left|\int_\lambda^{\lambda+\eta} y(x,\lambda')d\lambda'\right|^2 \right] = 1 \qquad (30\cdot 5)$$

for all values of λ in Ω_2. The Fourier coefficients c_n, $c(\lambda)$ are to be determined by the formulas

$$c_n = (\rho f(x), y_n(x)); \qquad c(\lambda) = \lim_{\eta \to 0} \eta^{-1}\left(\rho f(x), \int_\lambda^{\lambda+\eta} y(x,\lambda')d\lambda'\right). \qquad (30\cdot 6)$$

---

[1] If the "density function" ρ is a constant—the usual case—it is best to omit ρ in (30·5), (30·6), and (30·7). It then drops out of (30·8) and (30·9) automatically.

If $f(x)$ is absolutely integrable and $y(x,\lambda)$ has an upper bound, the latter formula can be replaced by the simpler one

$$c(\lambda) = (\rho f(x), y(x,\lambda)). \tag{30.7}$$

Weyl does not take up the question of completeness, but term-by-term integration gives the completeness relation

$$(\rho f, f) = \sum_{\text{all } n} |c_n|^2 + \int_{\Omega_2} |c(\lambda)|^2 d\lambda \tag{30.8}$$

for the class of functions $f(x)$ specified. Furthermore it is possible to extend this theorem to arbitrary piecewise continuous quadratically integrable functions $\psi(x)$ by the method cited in Sec. 25. Finally we can use the method of Sec. 22 [cf. Eq. (22.32) ff.] to derive the scalar-product relation

$$(\rho f, g) = \int_a^\infty \rho f g^* dx = \sum_{\text{all } n} c_n b_n^* + \int_{\Omega_2} c(\lambda) b^*(\lambda) d\lambda \tag{30.9}$$

for arbitrary piecewise continuous functions f, g, quadratically integrable with respect to $\rho(x)$ and having the Fourier coefficients c_n, $c(\lambda)$, and b_n, $b(\lambda)$, respectively.

*30c. Formal Treatment of Continuous Spectrum as the Limit of a Discrete Spectrum.—In this and succeeding sections we supplement the above summary of the results obtained by Weyl with a heuristic discussion intended to motivate conclusions whose origin must otherwise be veiled in an atmosphere of mystery. For simplicity we restrict our elementary treatment of the problem to the special case of the linear oscillator equation (cf. Secs. 18 and 19)

$$y'' + \kappa(E - V(x))y = 0. \qquad \kappa = \frac{8\pi^2\mu}{h^2} \tag{30.10}$$

Here E plays the part of λ in (30.1). We assume that $V(x)$ has a pole of the first or higher order at the left-hand boundary point $x = a$ and approaches a definite limit at $x = +\infty$ which we can identify with the zero level of the energy E.[1] $V(x)$ is assumed to be continuous at every interior point of the fundamental interval $a \leqslant x < +\infty$.

Under these circumstances we can satisfy the singular-point boundary conditions at $x = a$ for every E by choosing an integral curve $y(x)$ which has a node at $x = a$. Similarly we can satisfy the s.p.b.c. at $+\infty$ for every negative energy by choosing an integral curve for which

$$y(+\infty) = 0.$$

Solutions of (30.10) which satisfy the s.p.b.c. at either end of the funda-

---

[1] The argument would be sensibly unchanged by supposing that the left-hand boundary point is at $-\infty$ and that $\lim V(x) = +\infty$.

.mental interval have a finite number of zeros near that end point.
There are no positive energy solutions which satisfy the s.p.b.c. at $+\infty$,
or which are quadratically integrable over the entire fundamental
interval. An integral curve of (30·10) for $E > 0$ with a node at $x = a$
constitutes a type B eigenfunction. We designate by $y(x,E)$ a real
family of such eigenfunctions which is continuous in E. This family
will be uniquely defined when we have fixed on a definite scheme for
normalization.

The problem to be dealt with is that of finding the eigenvalues and
eigenfunctions of (30·10) for the infinite fundamental interval $a \leqslant x < +\infty$
and of setting up a scheme for expanding an arbitrary quadratically
integrable and piece-by-piece continuous $f(x)$ spread out over that interval
into a discrete-continuous linear combination of eigenfunctions. This
we designate as *problem α*. For comparison purposes it is convenient to
introduce a modified problem based on the same differential equation,
but using a finite fundamental interval $a \leqslant x \leqslant b$. This second problem
will be designated as *problem β*. We employ the s.p.b.c. at the left-hand
boundary point $x = a$ and give the problem a purely discrete spectrum
by requiring that all eigenfunctions shall vanish at the right-hand
boundary, $x = b$.

Every eigenfunction of β for a *positive* eigenvalue can be made to
coincide throughout the interval $a \leqslant x \leqslant b$ with the function $y(x,E)$ for
the same energy, provided that we choose a suitably modified normaliza-
tion for the former in the place of the usual one. Let $\epsilon_n(b)$ denote the
nth discrete eigenvalue of problem β and let $w_{nb}(x)$ denote the correspond-
ing eigenfunction. We adopt the modified normalization of the w's,
so that, for positive $\epsilon_n(b)$,

$$w_{nb}(x) = y[x, \epsilon_n(b)]; \qquad a \leqslant x \leqslant b$$
$$\qquad\quad = \text{undefined}. \qquad b < x \qquad \cdot \tag{30·11}$$

Hence it is not necessary for most purposes to distinguish between $w_{nb}(x)$
and $y[x, \epsilon_n(b)]$ when the energy is positive. If $\epsilon_n(b)$ is negative, $w_{nb}(x)$
does not coincide with an eigenfunction of α, although it is a solution of
its differential equation. Figure 11 shows a possible potential-energy
curve and illustrates qualitatively the nature of the eigenfunctions of
problem β.

As explained in Sec. 18, an increase in energy is always accompanied by
a decrease in the spacing of the nodes of the integral curves and *vice versa*.
Hence, if we consider only those integral curves of (30·10) which have a
node at $x = a$, a movement of the remaining nodes to the right must be
accompanied by a decrease in the energy. Since the last node of each
eigenfunction of problem β occurs at $x = b$, an increase in b must lower
all the energy eigenvalues. If these energies are plotted as ordinates

against b as abscissa, one obtains a set of monotonic decreasing curves as shown in Fig. 12.

Since the discrete eigenfunctions of problem α have nodes at infinity, it is clear that each class A solution of problem α (if any) is the limit,
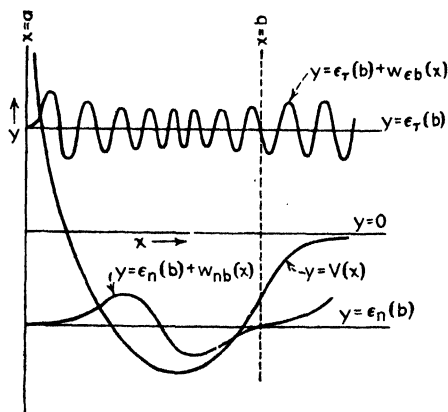
Fig. 11.—Illustrating problems α and β. Eigenfunctions for two energy levels of problem β are drawn in, using the corresponding lines of constant energy as axes of abscissas.

as b becomes infinite, of the corresponding solution of problem β. There may be an infinite number of discrete energy levels, a finite number, or none at all, in the spectrum of problem α depending on the form of $V(x)$. In the first mentioned case every energy level of problem β goes into a discrete level of problem α when b becomes infinite. Nevertheless,
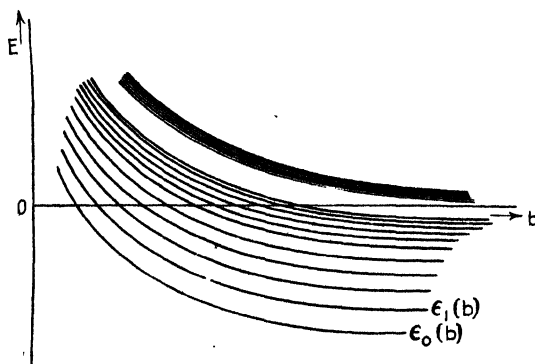
Fig. 12.—The energy levels of problem β as functions of b.

even in this case the limit of the complete spectrum of β, as b becomes infinite, is not a purely discrete spectrum, for, despite the fact that positive eigenvalues are constantly becoming negative with increasing b, the number of levels in any interval $0 < \epsilon < E$ increases without limit as $b \to \infty$. In other words the spacing of the levels in the neighborhood of any fixed positive energy value approaches zero as $b \to \infty$.

*30d. The Spacing of Energy Levels in Problems β and α.—The above proposition is conveniently proved by setting up an explicit formula for the spacing of the energy levels in problem β. Such a formula is derived in Appendix H with the following result:

$$\Delta\epsilon(n,b) \equiv \epsilon_{n+1}(b) - \epsilon_n(b) = \frac{l_n(b)}{\kappa} \frac{y_x[b', \epsilon_n(b')]^2}{\int_a^{b'} y[x, \epsilon_n(b)]^2 dx}. \tag{30.12}$$

Here $l_n(b)$ is the distance from b to the preceding node of $y[x, \epsilon_n(b)]$, $y_x[b', \epsilon_n(b')]$ is the derivative with respect to x of $y[x, \epsilon_n(b')]$ at $x = b'$, and b' is a suitably chosen real number in the half wave-length interval

$$b - l_n(b) < b' < b.$$

It is further proved in Appendix H that any integral curve of (30.10) for a positive value of ϵ, say $y(x,E)$, can be fitted to a suitably chosen Brillouin-Wentzel-Kramers approximation function [*cf.* Sec. 21*a*, Eq. (21.7)]

$$u(x,E) = A p^{-\frac{1}{2}} \cos\left[\frac{2\pi}{h}\int_{x_0}^{x} p(x,E)dx + \gamma\right], \qquad p^2 = 2\mu(E - V) \tag{30.13}$$

so that, for $x - x_0 > 0$,

$$|y(x,E) - u(x,E)| < \frac{M}{x^2}$$

where M is a positive number which becomes infinite when E approaches zero.

Let us now allow n and b to increase together from any pair of initial values n_0, b_0 to infinity in such a manner that $\epsilon_n(b) = E$, independent of n.
Then[1]

$$\lim_{n,b \to \infty} l_n(b) = \frac{h}{2\sqrt{2\mu E}}. \tag{30.14}$$

It follows from (30.13) that, to a close approximation for large values of b,

$$y_x[b', \epsilon_n(b')]^2 = \frac{4\pi^2 A^2}{h^2}\sqrt{2\mu[\epsilon_n(b') - V(b')]}. \tag{30.15}$$

We conclude that the numerator of the right-hand member of Eq. (30.12) remains finite, as n and b become infinite. The denominator becomes infinite, however, so that

$$\lim_{n,b \to \infty} \Delta\epsilon(n,b) = 0. \tag{30.16}$$

---

[1] This can be proved either by the approximation formula (30.13) or by Sturm's fundamental oscillation theorem (*cf.* Ince, *Ordinary Differential Equations*, p. 224).

Hence *the density of the energy levels of problem β in the neighborhood of any positive energy E becomes infinite with b*, and the complete spectrum of the β problem must go over into the sum of a discrete spectrum of negative energies and a continuous spectrum of positive energies.

It follows from the definition of b' that

$$\epsilon_{n-1}[b - l_n(b)] < \epsilon_n[b'] < \epsilon_n[b - l_n(b)].$$

In view of (30·16) we infer that $\epsilon_n(b')$ approaches $\epsilon_n(b)$ as b and n become infinite. Consequently (30·12), (30·14), and (30·15) yield

$$\lim_{n,b \to \infty} \Delta\epsilon(n,b) \int_a^b y[x, \epsilon_n(b)]^2 dx = \frac{2\pi^2 A^2}{\kappa h} = \frac{A^2 h}{4\mu}. \tag{30·17}$$

In view of (30·17) it will simplify matters if we agree to *normalize the family of type B eigenfunctions y(x,E) by the requirement that for all positive energies the coefficient A of (30·13) shall have the value $2\sqrt{\mu/h}$.* The right-hand member of (30·17) then reduces to unity, and we can replace $\int_a^b y(x, \epsilon_n(b))^2 dx$ by $(\Delta\epsilon(n,b))^{-1}$ with an error which approaches zero when b and n become infinite together in such a way that $\epsilon_n(b)$ remains constant.

*30e. The Eigendifferentials.—Consider next the sum $S_{(E_\tau)}^{(E_\tau+\eta_\tau)}(x,b)$ defined by the equation

$$S_{(E_\tau)}^{(E_\tau+\eta_\tau)}(x,b) = \sum_{\substack{\epsilon_n(b) < E_\tau+\eta_\tau \\ \epsilon_n(b) > E_\tau}} w_{nb}(x)\Delta\epsilon(n,b). \tag{30·18}$$

For any fixed values of x, E_τ, η_τ, we obtain an eigendifferential from the above expression by allowing b to become infinite. Thus

$$\Delta_\tau y = \int_{E_\tau}^{E_\tau+\eta_\tau} y(x,E')dE' = \lim_{b \to \infty} S_{(E_\tau)}^{(E_\tau+\eta_\tau)}(x,b). \tag{30·19}$$

Moreover, due to the orthogonality of the functions $w_{nb}(x)$,

$$\|S_{(E_\tau)}^{(E_\tau+\eta_\tau)}(x,b)\|^2 \equiv \int_a^b [S_{(E_\tau)}^{(E_\tau+\eta_\tau)}(x,b)]^2 dx = \sum_{\substack{\epsilon_n(b) < E_\tau+\eta_\tau \\ \epsilon_n(b) > E_\tau}} \int_a^b [w_{nb}(x)]^2 dx [\Delta\epsilon(n,b)]^2.$$

Substituting $[\Delta\epsilon(n,b)]^{-1}$ for $\int_a^b [w_{nb}(x)]^2 dx$ and taking the limit as b becomes infinite, we obtain

$$\lim_{b \to \infty} \int_a^b [S_{(E_\tau)}^{(E_\tau+\eta_\tau)}(x,b)]^2 dx = \lim_{b \to \infty} \sum_{E_\tau}^{E_\tau+\eta_\tau} \Delta\epsilon(n,b) = \eta_\tau.$$

Finally, by interchanging the order of two limiting processes, we obtain

$$\|\Delta_\tau y\|^2 = (\Delta_\tau y, \Delta_\tau y) = \eta_\tau. \tag{30·20}$$

This formula indicates[1] the quadratic integrability of $\Delta_\tau y$ and *is equivalent to the normalization condition* (30·5). As we have derived it on the assumption that the amplitude coefficient at infinity, *i.e.*, A, has the value $2\sqrt{\mu/h}$, we see that (30·20) must be equivalent to giving A the value in question.

The *orthogonality of the eigendifferentials* $\Delta_\tau y$ *for two non-overlapping intervals* $E_\tau < E < E_\tau + \eta_\tau$, $E_{\tau'} < E < E_{\tau'} + \eta_{\tau'}$ can be deduced as follows. If we substitute for $\Delta_\tau y$ in (30·20) the sum $\Delta_\tau y + \Delta_{\tau'} y$ we must evidently set the right-hand member equal to $\eta_\tau + \eta_{\tau'}$. Hence

$$(\Delta_\tau y, \Delta_\tau y) + (\Delta_{\tau'} y, \Delta_\tau y) + (\Delta_\tau y, \Delta_{\tau'} y) + (\Delta_{\tau'} y, \Delta_{\tau'} y) = \eta_\tau + \eta_{\tau'}.$$

Then applying (30·20) to the individual eigendifferentials $\Delta_\tau y$, $\Delta_{\tau'} y$, and remembering that $\Delta_{\tau'} y$ and $\Delta_\tau y$ are real, we obtain

$$(\Delta_\tau y, \Delta_\tau y) = 0.$$

The derivative of $y(x,E), (E > 0)$, with respect to x behaves in the same way at infinity as $y(x,E)$ itself. This follows from the asymptotic B. W. K. approximation. Hence $\dfrac{d}{dx}\Delta_\tau y$ is quadratically integrable with respect to x and must approach zero at infinity. We conclude that $\Delta_\tau y$ satisfies the s.p.b.c. at both ends of the fundamental interval. It is in fact a type A function and belongs to a class of functions with respect to which the operator $H = -\dfrac{1}{\kappa}\dfrac{d^2}{dx^2} + V(x)\times$ is Hermitian (*cf.* Sec. 32d, p. 202).

Let us now consider the function

$$S(x) = \int_{E'}^{E''} \varphi(E)y(x,E)dE, \tag{30·21}$$

where φ is a piece-by-piece continuous function of E, and both E' and E'' are positive. We can approximate this function for any fixed x by a sum of products of eigendifferentials $\Delta_\tau y$ each multiplied by the corresponding $\varphi(E_\tau)$. Forming the scalar product of two such approximations and taking the limit as the range of each eigendifferential becomes zero, we obtain

$$\|S\|^2 \equiv (S(x),S(x)) = \lim_{\eta_\tau \to 0}\left[\sum_{E' < E_\tau}^{E_\tau < E''} |\varphi(E_\tau)|^2 \eta_\tau\right] = \int_{E'}^{E''}|\varphi(E)|^2 dE. \tag{30·22}$$

This is the equivalent of Plancherel's theorem (9·6) for functions of the form of $S(x)$.

---

[1] Of course our procedure has not been rigorous, owing to the interchange of the order of two limiting processes. A rigorous proof of the quadratic integrability of $\Delta_\tau y$ can be derived from the B. W. K. approximation (30·13).

*30f. Passage from the Completeness Theorem for Problem β to That for Problem α.—If the functions $w_{nb}(x)$ are not normalized according to the usual rule for discrete eigenfunctions the completeness theorem for problem β and quadratically integrable functions $f(x)$ takes the form

$$\int_a^b |f|^2 dx = \sum_0^\infty |c_{nb}|^2 \int_a^b [w_{nb}(x)]^2 dx, \qquad (30\cdot23)$$

where c_{nb} is the Fourier coefficient

$$c_{nb} = \frac{\int_a^b f(x) w_{nb}(x) dx}{\int_a^b [w_{nb}(x)]^2 dx}.$$

Let us assume that the functions $w_{nb}(x)$, whose energies are negative, are normalized in the usual manner for discrete eigenfunctions, while those for positive energies are normalized according to (30·11). We insert the value $\int_a^b f(x) w_{nb}(x) dx \Big/ \left(\int_a^b [w_{nb}(x)]^2 dx \right)$ for c_{nb} when $\epsilon_n(b) > 0$. Then (30·23) becomes

$$\sum_{\epsilon_n(b) < 0} |c_{nb}|^2 + \sum_{\epsilon_n(b) \geqslant 0} \frac{\left| \int_a^b f(x) w_{nb}(x) dx \right|^2}{\int_a^b [w_{nb}(x)]^2 dx} = \int_a^b |f|^2 dx. \qquad (30\cdot24)$$

Let us now assume that $f(x)$ is normalized to unity in the interval

$$a < x < +\infty$$

and that it is absolutely, as well as quadratically, integrable in this interval. The latter condition together with the existence of an upper bound for the type B eigenfunctions $y(x,E)$ permits us to define the expansion coefficient $c(E)$ by the formula

$$c(E) \equiv \int_a^\infty f(x) y(x,E) dx \qquad (30\cdot25)$$

in the assurance that the integral converges, since

$$\left| \int_a^\infty f(x) y(x,E) dx \right| \leqslant \int_a^\infty |f| |y(x,E)| dx$$

$$\leqslant |y(x,E)|_{\max} \int_a^\infty |f| dx.$$

In laying down the definition (30·25) we deliberately omit the factor $\rho(x) = \kappa$ of (30·7) in accordance with the suggestion of footnote 1, p. 164.

In view of (30·17) and the normalization of A to the value $2\sqrt{\mu/h}$, we can replace $\int_a^b [w_{nb}(x)]^2 dx$ by $\Delta\epsilon(n,b)^{-1}$ with an error which approaches zero as b becomes infinite. This yields

$$\sum_{E_n < 0} |c_n|^2 + \lim_{b \to \infty} \sum_{\epsilon_n(b) \geq 0} \left| \int_a^b f(x) y[x, \, \epsilon_n(b)] dx \right|^2 \Delta \epsilon(n,b) = \int_a^\infty |f|^2 dx = 1. \,\, (30 \cdot 26)$$

In view of (30·25) the above equation goes over (formally) into

$$\sum_{\text{all } n} |c_n|^2 + \int_0^\infty |c(E)|^2 dE = \int_a^\infty |f|^2 dx = 1. \tag{30·27}$$

As a corollary on (30·27) and (30·22) we infer that

$$\left\| f - \sum_{\text{all } n} c_n y_n(x) - \int_0^\infty c(E) y(x,E) dE \right\|^2 = 0. \tag{30·28}$$

Finally, if $g(x)$ is a second function of x quadratically integrable in $a < x < \infty$, and if $b_n, b(E)$ are its Fourier coefficients,

$$(f,g) = \sum_{\text{all } n} c_n b_n{}^* + \int_0^\infty c(E) b^*(E) dE. \tag{30·29}$$

In view of these relations we say that the totality of the type A and type B eigenfunctions of problem α forms a *complete system*.

An exact type A solution of (30·10) reduces $\|(H - E)y\|$ to zero if we identify H with $-\dfrac{1}{\kappa} \dfrac{d^2}{dx^2} + V(x)$. Hence any type A function can be said to form a good approximate solution of (30·10) if it makes

$$\frac{((H - E)y, (H - E)y)}{(y,y)} \ll 1$$

for some value of E. From this point of view it is easy to use (30·29) to prove that the general eigendifferential $\Delta y = \int_E^{E+\eta} y(x,E') dE'$ is actually a very good approximate type A solution of (30·10) when η is small. Thus

$$\frac{((H - \epsilon)\Delta y, (H - \epsilon)\Delta y)}{(\Delta y, \Delta y)} = \frac{1}{\eta} [(H\Delta y, H\Delta y) - 2\epsilon(H\Delta y, \Delta y) + \epsilon^2(\Delta y, \Delta y)]$$

$$= \frac{1}{\eta} \Big[\int_E^{E+\eta} E'^2 dE' - 2\epsilon \int_E^{E+\eta} E' dE' + \epsilon^2 \eta \Big].$$

When ϵ is given the value $E + \frac{1}{2}\eta$, the above reduces to

$$\frac{((H - \epsilon)\Delta y, (H - \epsilon)\Delta y)}{(\Delta y, \Delta y)} = \frac{\eta^2}{12}.$$

Hence we can make Δy as good an approximate type A eigenfunction as we please by a proper choice of η in spite of the fact that $\lim_{\eta \to 0} \Delta y$ is not itself a type A function.

Equation (30·4) can be set up formally for our special case by a procedure similar to that adopted for the completeness formula (30·29).

It is, of course, to be expected that (30·27), (30·28), and (30·29) will hold for all quadratically integrable functions f, g, whereas (30·4) cannot be assumed valid outside the restricted class of function $f(x)$ admitted by the Weyl theory.

30g. The Fourier Integral Formulas a Special Case.—As a first simple illustration of the theory of continuous spectra let us consider the equation of the linear oscillator (30·10) in the interval $0 < x < \infty$ with $V(x)$ set equal to zero and the boundary condition $y(0) = 0$. This may be interpreted as the problem of a free particle in one dimension with an abrupt infinite potential barrier at the origin. The boundary condition is really a special case of the s.p.b.c. where the end point is not singular. The eigenfunctions are multiples of $\sin [(\kappa E)^{\frac{1}{2}}x]$ and, when normalized in accordance with the rule following Eq. (30·17) [or by Eq. (30·20)], take the definite form

$$y(x, E) = 2\sqrt{\frac{\mu}{h}}(2\mu E)^{-\frac{1}{4}}\sin[(\kappa E)^{\frac{1}{2}}x]. \qquad (30\cdot30)$$

We take $f(x)$ to be an arbitrary continuous absolutely and quadratically integrable function defined in the interval $0 < x < \infty$. The expansion (30·4) then becomes

$$f(x) = 2\sqrt{\frac{\mu}{h}}\int_0^\infty c(E)(2\mu E)^{-\frac{1}{4}}\sin[(\kappa E)^{\frac{1}{2}}x]dE,$$

where $c(E)$, in accordance with (30·25), has the value

$$c(E) = 2\sqrt{\frac{\mu}{h}}(2\mu E)^{-\frac{1}{4}}\int_0^\infty f(\xi)\sin[(\kappa E)^{\frac{1}{2}}\xi]d\xi.$$

Replacing κE by $4\pi^2\sigma^2$, we obtain the symmetric Fourier integral formulas[1]

$$\left.\begin{array}{l} f(x) = 2\displaystyle\int_0^\infty G(\sigma)\sin 2\pi x\sigma\, d\sigma, \\[2mm] G(\sigma) = 2\displaystyle\int_0^\infty f(x)\sin 2\pi x\sigma\, dx. \end{array}\right\} \qquad (30\cdot31)$$

In case we abandon the boundary condition at the origin and suppose our free particle to range from $-\infty$ to $+\infty$, it is evident that we shall have to use a generalization of the theory of continuous spectra given above. The phase of the eigenfunctions is now completely indeterminate and $f(x)$ must be expanded in terms of the two linearly independent eigenfunctions $\sin [(\kappa E)^{\frac{1}{2}}x]$ and $\cos [(\kappa E)^{\frac{1}{2}}x]$. *In this case the eigenfunctions of the continuous spectrum are degenerate.* Using appropriate normalization and determining the expansion coefficients in the same way we obtain the Fourier integral formulas of Sec. 9.

[1] These formulas hold if $f(-x) = -f(x)$, or, if $f(0) = 0$ and $f(x)$ is undefined for negative values of x. They are readily derived from the form of the Fourier integral theorem used in Sec. 9.

30h. Normal Packet Functions in One Dimension.—We have defined normal packet functions (Sec. 10) as quadratically integrable non-monochromatic solutions of Schrödinger's second wave equation (5·10). Very general solutions of this kind are obtained by taking linear combinations of eigenfunctions of the first Schrödinger equation, each multiplied by an appropriate time factor. Thus, if H is the one-dimensional Hamiltonian operator $V(x) - \dfrac{1}{\kappa}\dfrac{d^2}{dx^2}$ for (30·10), and if the series-integral

$$\Psi(x,t) \equiv \sum_n c_n \psi_n(x) e^{-\frac{2\pi i}{h} E_n t} + \int_0^\infty c(E)\psi(x,E) e^{-\frac{2\pi i}{h} E t} dE \qquad (30·32)$$

is uniformly convergent, term-by-term application of the operator $H + \dfrac{h}{2\pi i}\dfrac{\partial}{\partial t}$ suffices to prove that it represents a solution of the second Schrödinger equation. If the function $\Psi(x,t)$ defined by the series is quadratically integrable, it is a normal packet function. In Sec. 5 it was proved that every complete wave function $\Psi(x,t)$ is determined by its differential equation and its form at some arbitrary time, say $t = 0$. It follows that the most general normal packet function for (30·10) is obtainable by fitting the initial condition $\Psi(x,0) = f(x)$, where $f(x)$ is given the most general form for a type A function.

But we know from the Weyl theory that if $\varphi(x)$ is any function which is quadratically integrable, together with its transform $H\varphi$, over the interval $a < x < \infty$, it is expansible into the uniformly convergent discrete-continuous linear combination of eigenfunctions of H,

$$\varphi(x) = \sum_n d_n \psi_n(x) + \int_0^\infty d(E)\psi(x,E)dE, \qquad (30·33)$$

We have only to identify the expansion coefficients in (30·32) with those in (30·33) in order to get a normal packet function which reduces to $\varphi(x)$ at $t = 0$.

It is not immediately evident that every type A function $f(x)$ necessarily has a quadratically integrable transform by H, but in Sec. 32 this requirement will be added as a new restriction on the class of physically admissible functions. From this point of view we can then say that every physically admissible solution of the one-dimensional equation $H\Psi = -\dfrac{h}{2\pi i}\dfrac{\partial \Psi}{\partial t}$ satisfies the conditions imposed on $\varphi(x)$ and is expressible in the form (30·32).

30i. Normal Packet Functions for the Two-particle Problem.—As a further illustration of the theory of continuous spectra let us return to the two-particle problem of Sec. 28. The equation for the radial eigenfunctions $\Re$, viz., (28·19), is of the same type as Eq. (30·10). Hence

an arbitrary function of the radius which conforms to the Weyl conditions (p. 164) is capable of expansion in the form

$$u(r) = \sum_{n=l+1}^{\infty} c_n \mathfrak{R}_{nl}(r) + \int_{V(\infty)}^{\infty} c(E) \mathfrak{R}_l(r,E) dE, \qquad (30\cdot34)$$

where $\mathfrak{R}_{nl}$ and $\mathfrak{R}_l(r,E)$ are normalized type A and type B eigenfunctions of Eq. (28·19), respectively. It follows that *an essentially arbitrary normal packet function for the two-particle problem can be expressed in the form*

$$\Psi(r,\theta,\varphi,t) = \sum_{m=-\infty}^{+\infty} \sum_{l=|m|}^{\infty} \Theta_{lm} \Phi_m \left[\sum_{n=l+1}^{\infty} c_{nlm} R_{nl} e^{-\frac{2\pi i}{h} E_{nl} t} \right.$$
$$\left. + \int_{V(\infty)}^{\infty} c_{lm}(E) R_l(r,E) e^{-\frac{2\pi i}{h} Et} dE \right]. \quad (30\cdot35)$$

To prove the above proposition we observe first of all that the above expression is a solution of the appropriate second Schrödinger wave equation

$$H\Psi \equiv -\frac{h^2}{8\pi^2\mu} \left(\nabla^2 \Psi - \frac{8\pi^2\mu}{h^2} V\Psi \right) = -\frac{h}{2\pi i} \frac{\partial \Psi}{\partial t}. \qquad (30\cdot36)$$

In the second place it can be fitted to arbitrary physically admissible initial conditions. Let

$$F(x,y,z) \equiv f(r,\theta,\varphi),$$

for example, denote an arbitrary type A function of the Cartesian coordinates with a quadratically integrable transform HF. Following an argument which parallels that used at the close of Sec. 29 we expand $f(r,\theta,\varphi)$ in the uniformly convergent series of spherical harmonics

$$f(r,\theta,\varphi) = \sum_{m=-\infty}^{+\infty} \sum_{l=|m|}^{\infty} G_{lm}(r) \Theta_{lm}(\theta) \Phi_m(\varphi). \qquad (30\cdot37)$$

The coefficients G_{lm} are undetermined continuous functions of r.

We can also expand Hf into a series of spherical harmonics

$$Hf = \sum_{m=-\infty}^{+\infty} \sum_{l=|m|}^{\infty} F_{lm}(r) Y_{lm}(\theta,\varphi). \qquad Y_{lm} = \Theta_{lm}(\theta)\Phi_m(\varphi) \quad (30\cdot38)$$

Let us assume for the moment our right to apply the operator H term by term to the expansion (30·37). Introducing the definitions

$$\left. \begin{array}{c} \mathfrak{g}_{lm} \equiv r G_{lm}, \\[2mm] \Lambda_l \equiv -\frac{h^2}{8\pi^2\mu} \left[\frac{d^2}{dr^2} - \frac{8\pi^2\mu}{h^2} V(r) - \frac{l(l+1)}{r^2} \right], \end{array} \right\} \qquad (30\cdot39)$$

we readily derive

$$Hf = \sum_m \sum_l HG_{lm}Y_{lm} = \sum_m \sum_l r^{-1}(\Lambda_l \mathcal{G}_{lm})Y_{lm}. \qquad (30{\cdot}40)$$

Hence

$$F_{lm}(r) = r^{-1}\Lambda_l \mathcal{G}_{lm}.$$

In Appendix I it is proved that the term-by-term application of the Hamiltonian operator H is valid and that the integrals

$$\int_0^\infty r^2|G_{lm}|^2 dr = \int_0^\infty |\mathcal{G}_{lm}|^2 dr; \qquad \int_0^\infty r^2|F_{lm}|^2 dr = \int^\infty |\Lambda_l \mathcal{G}_{lm}|^2 dr$$

are convergent. Thus $\mathcal{G}_{lm}(r)$ conforms to the conditions imposed on the function f of Eq. (30·4) by the Weyl theory and can be expanded into a series-integral combination of normal eigenfunctions and type B eigenfunctions for Eq. (28·19). We write

$$\mathcal{G}_{lm} = \sum_{n=l+1}^\infty c_{nlm}\mathcal{R}_{nl} + \int_{V(\infty)}^\infty c_{lm}(E)\mathcal{R}_l(r,E)dE, \qquad (30{\cdot}41)$$

which yields

$$f(r,\theta,\varphi) = \sum_{m=-\infty}^{+\infty} \sum_{l=|m|}^\infty \Theta_{lm}\Phi_m\left[\sum_{n=l+1}^\infty c_{nlm}R_{nl} + \int_{V(\infty)}^\infty c_{lm}(E)R_l(r,E)dE\right]. \tag{30·42}$$

Comparing Eqs. (30·35) with (30·42), we see that $\Psi(r,\theta,\varphi,t)$ can be fitted to the arbitrary physically admissible initial function $f(r,\theta,\varphi)$ with quadratically integrable transform Hf by a proper choice of the coefficients c_{nlm}, $c_{lm}(E)$. This proves the theorem.

If the functions $\Phi_m(\varphi), \Theta_{lm}(\theta), R_{nl}(r)$ and $R_l(r,E)$ are normalized, the formulas for the coefficients are

$$c_{nlm} = (\Psi, \Psi_{nlm}) = \iiint_\infty \Psi R_{nl}\Theta_{lm}\Phi_m{}^* e^{\frac{2\pi i}{h}E_{nl}} r^2 \sin\theta \, dr d\theta d\varphi, \qquad (30{\cdot}43)$$

$$\left.\begin{array}{l} c_{nl}(E) = \lim_{E' \to E} \dfrac{1}{E'' - E}(\Psi, \Delta\Psi_{lm}[E]), \\[2mm] \Delta\Psi_{lm}[E] = \displaystyle\int_E^{E'} \psi_{lm}(E'') e^{-\frac{2\pi i}{h}E''t} dE''. \end{array}\right\} \qquad (30{\cdot}44)$$

The completeness of the system of eigenfunctions $\psi_{nlm}, \psi_{lm}(E)$ in the sense of Eq. (30·8) is a corollary on the expansion theorem.

*30j. Normalization of the Class B Radial Eigenfunctions for the Hydrogenic Atom.—In the special case of the hydrogenic-atom problem, where $V(r)$ represents a Coulomb potential, it is possible to give explicit formulas for the radial eigenfunctions of the continuous spectrum in

terms of complex integrals.[1] Similar formulas are available for the discrete eigenfunctions but will not be used here. The substitution

$$\Re_l(\rho,\epsilon) = \rho^{l+1}S_l(\rho,\epsilon) \qquad (30\cdot45)$$

in Eq. (29·2) yields the Laplace-type differential equation

$$\frac{d^2S_l}{d\rho^2} + \frac{2(l+1)}{\rho}\frac{dS_l}{d\rho} + \left(\epsilon + \frac{2}{\rho}\right)S_l = 0, \qquad (30\cdot46)$$

which can be reduced to the first order by Laplace's transformation. It has positive energy solutions of the form[2]

$$S_l(\rho,\epsilon) = a_r\int_{-i\sqrt{\epsilon}}^{+i\sqrt{\epsilon}} e^{\rho z}(z - i\sqrt{\epsilon})^{l-i/\sqrt{\epsilon}}(z + i\sqrt{\epsilon})^{l+i/\sqrt{\epsilon}}\,dz, \quad (30\cdot47)$$

where a_r is a normalization factor whose value is to be determined. Useful expansions for small and large values of ρ can be derived from this integral. For large positive values of ρ Fues (see footnote) gives the asymptotic expression

$$S_l(\rho,\epsilon) \cong -2ia_r\rho^{-(l+1)}[T \cos \delta(\rho,\epsilon) + S \sin \delta(\rho,\epsilon)], \qquad (30\cdot48)$$

where

$$\left.\begin{array}{l} \delta(\rho,\epsilon) = \rho\sqrt{\epsilon} + \dfrac{1}{\sqrt{\epsilon}} \log \rho; \\[2mm] S + iT = (2\sqrt{\epsilon})^{l+\frac{i}{\sqrt{\epsilon}}} e^{-\frac{3\pi}{2\sqrt{\epsilon}} - \frac{i\pi}{2}(l+2)}\Gamma\left(l + 1 - \dfrac{i}{\sqrt{\epsilon}}\right). \end{array}\right\} \qquad (30\cdot49)$$

Elementary manipulation yields

$$\Re_l(\rho,\epsilon) \cong -2ia_r(2\sqrt{\epsilon})^l e^{-\frac{3\pi}{2\sqrt{\epsilon}}}\left|\Gamma\left(l + 1 - \frac{i}{\sqrt{\epsilon}}\right)\right| \cos\left[\delta(\rho,\epsilon) + \beta(l,\epsilon)\right].$$

The value of the phase constant $\beta(l,\epsilon)$ is not worked out by Fues, but may be computed from formulas which he gives if desired.

According to the normalization scheme of (30·20) we should convert $\Re_l(\rho,\epsilon)$ into an explicit function of r and E, and then choose a_r to meet the condition[3]

$$\int_0^\infty \left|\int_E^{E+\eta} \Re_l\left(\frac{rZ}{a_0}, \frac{a_0^2\kappa E}{Z^2}\right)dE\right|^2 dr = \eta.$$

[1] *Cf.* E. SCHRÖDINGER, *Ann. d. Physik* **79**, 361 (1926).

[2] *Cf.* L. SCHLESINGER, *Einführung in die Theorie der gewöhnlichen Differential-gleichungen*, 3d ed., Kap. 8, Berlin, 1922; E. FUES, *Ann. d. Physik* **80**, 367 (1926), **81**, 281 (1926).

[3] Of course we are at liberty to simplify our formulas by *revising* the method of normalization, using ρ as the variable of integration in forming scalar products and substituting ϵ for E in (30·20) and in all formulas depending on the normalization scheme. The details can be left to the reader.

This is equivalent to writing $\mathcal{R}_l(\rho,\epsilon)$ in the form (30·13) with A set equal to $2\sqrt{\mu/\hbar}$. In other words we must choose a_r so that, for large values of r,

$$\mathcal{R}_l\left(\frac{rZ}{a_0}, \frac{a_0^2\kappa E}{Z^2}\right) = \frac{2}{(2\pi e^2 Z)^{1/2}}\left[\epsilon + \frac{2}{\rho} - \frac{l(l+1)}{\rho^2}\right]^{-1/4}$$

$$\times \cos\left[\int_{\rho_0}^{\rho}\left[\epsilon + \frac{2}{u} - \frac{l(l+1)}{u^2}\right]^{1/2} du + \beta'(l,\epsilon)\right].$$

This means that

$$a_r = \frac{i}{(2\pi e^2 Z)^{1/2}}\frac{e^{\frac{3\pi}{2\sqrt{\epsilon}}}\,c^{-\left(\frac{l}{2}+\frac{1}{4}\right)}}{2^l|\Gamma(l+1-i\epsilon^{-1/2})|}. \tag{30·50}$$

To obtain the complete expression for the normalized eigenfunctions of the continuous spectrum we combine Eqs. (30·45), (30·47), and (30·50), replace ρ by rZ/a_0 and multiply in the normalized expressions for Θ_{lm} and Φ_m. Thus

$$\psi_{lm}(r,\theta,\varphi,\epsilon) = r^{-1}\mathcal{R}_l\left(\frac{rZ}{a_0}, \epsilon\right)\Theta_{lm}(\theta)\Phi_m(\varphi)$$

$$= \frac{\left(\frac{2Z}{e^2}\right)^{1/2}e^{\frac{3\pi}{2\sqrt{\epsilon}}}}{2\pi a_0(2\sqrt{\epsilon})^{l+1/2}|\Gamma(l+1-i\epsilon^{-1/2})|}\sqrt{\frac{(2l+1)(l-|m|)!}{2(l+|m|)!}}\left(\frac{rZ}{a_0}\right)^l$$

$$\times P_{l|m|}(\cos\theta)e^{im\varphi}i\int_{-i\sqrt{\epsilon}}^{+i\sqrt{\epsilon}}e^{\frac{rzZ}{a_0}}(z-i\sqrt{\epsilon})^{l-\frac{i}{\sqrt{\epsilon}}}(z+i\sqrt{\epsilon})^{l+\frac{i}{\sqrt{\epsilon}}}dz. \tag{30·51}$$

The absolute value of the gamma function $\Gamma\left(l+1-\dfrac{i}{\sqrt{\epsilon}}\right)$ can be calculated by means of an infinite product given by Fues (*loc. cit.*, p. 303).

31. WEAK QUANTIZATION. THEORY OF RADIOACTIVE EMISSION OF ALPHA PARTICLES

31a. Weak Quantization in General.—It will be of interest to consider at this point the application of our theory of continuous spectra to problems involving imperfect, or weak, quantization.

In the Bohr theory, it will be remembered, a quantized, or stationary, state was one involving a multiply periodic classical motion and a discrete energy level determined by the Sommerfeld quantum conditions. It was early recognized by Bohr that where an approximate theory of an atomic system gives such quantized states with sharply defined energy levels, a more exact theory may show that the classical motions are only approximately of the multiply periodic type and that in consequence

the quantum conditions are not really applicable after all. Bohr had no scheme for adapting quantum properties to motions of other types than the multiply periodic one. Hence he had always to begin work by replacing the complete classical theory of the system under consideration by an idealized one having multiply periodic motions. It was necessary, for example, to discard radiation forces and interatomic collisions in order to set up a model of the hydrogen atom to which the quantum conditions would be applicable. The deficiencies of his theory were frankly faced by Bohr and he foresaw at an early date the probability that in cases where the classical motion is only approximately of the multiply periodic type a more complete theory would replace the sharply defined energy levels of his own theoretical methods with a narrow range of preferred energy values.[1]

Bohr's expectation has been fulfilled by the development of the quantum-mechanical theory of states of imperfect quantization. Both theoretically and experimentally we find states intermediate in character between the sharply defined stationary states of the typical discrete energy level and the communistic array of qualitatively similar states of a free electron. These intermediate states will hereafter be referred to as *weakly quantized* on account of their energy uncertainty. Contrary to what one might suppose from the above introductory remarks, they do not occur whenever a classical analysis would suggest an aperiodic motion, nor are they always excluded in cases in which classical theory does give periodic, or multiply periodic, motions. They are described by wave packets built up from continuous-spectrum eigenfunctions which approximate the properties of discrete eigenfunctions. Interpreting states with discrete- and continuous-spectrum eigenfunctions as the analogues of the classical multiply periodic and aperiodic motions, the parallelism between the quantum-mechanical situation and that previsaged by the Bohr theory becomes apparent.

31b. A Model for Alpha Particle Disintegration.—As an example we shall consider here the imperfectly quantized nuclear energy levels of the Gamow-Gurney-Condon theory of alpha-disintegration.[2] This theory is built on a simplified model which replaces the actual nucleus by a single alpha particle moving in a central force field. The model is similar in character to those used in dealing with optical atomic spectra, diatomic molecular spectra, and collisions between electrons and atoms, as two-particle problems. In the case under consideration the field is assumed to be of the Coulomb type down to a radius of the order of

---

[1] *Cf.* N. BOHR, *Zeits. f. Physik* **13,** 117 (1923); W. PAULI, JR., in GEIGER and SCHEEL'S *Handbuch der Physik,* vol. XXIII, p. 68, Berlin, 1926.

[2] *Cf.* G. GAMOW, *Zeits. f. Physik* **51,** 204 (1928); R. W. GURNEY and E. U. CONDON, *Phys. Rev.* **33,** 127 (1929); M. BORN, *Zeits. f. Physik* **58,** 306 (1929); H. CASIMIR, *Physica* **1,** 193 (1934).

10^{-12} cm., beyond which it passes through a maximum and then falls off rapidly to take on a minimum value at the origin. The potential hole at the center is analogous to the crater of a volcano, and one would actually get a volcano-shaped model if he plotted the potential energy as ordinate over a two-dimensional section of the field about the nucleus. The assumption of a Coulomb field outside the rim of the crater is well grounded on scattering experiments and the theory of atomic spectra. The crater itself is needed to provide the weakly quantized initial states from which the radioactive disintegration proceeds. If we wish to use the language of a less schematic theory we speak of a radioactive atom as a system of protons and neutrons, or alpha particles, neutrons, and protons in a weakly quantized state.
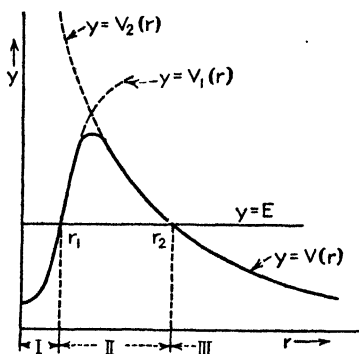
Fig. 13.—The volcano model for alpha-particle disintegration.

From a classical point of view this model yields no quasi-periodic motions. Alpha particles inside the crater have periodic or multiply periodic motions if their energy is less than V_{max}, while alpha particles outside the crater, or having an energy greater than V_{max}, have purely aperiodic motions. From the standpoint of quantum mechanics, however, the situation is quite different. Owing to the "tunnel effect" discussed in Sec. 21 the wave function of an alpha particle initially inside the crater, but having an energy greater than $V(\infty)$ and less than V_{max}, will in time leak through the crater wall. It follows that, as time goes on, the probability that the particle is inside the crater will decrease from its unit initial value, while the probability that it is outside increases correspondingly. Thus an assemblage of similar models described by such a "leaking" wave function would be observed to disintegrate spontaneously after the manner of real radioactive nuclei.

It is necessary to distinguish between those cases in which the crater is deep and wide enough to permit the existence of one or more discrete negative energy levels, and those in which the crater is too shallow for negative energies. In the *former case* there will be a lowest discrete energy level correlated with a nonradioactive normal state of the nuclear model. Only excited states of the nucleus will disintegrate spontaneously, for only excited states with positive energies will have a finite potential barrier to tunnel under—the tunnel must be at constant energy, *i.e.*, horizontal on such a graph as Fig. 13. In the *latter case*, however, *any* initial wave function representing a state in which the alpha particle starts out inside the crater will necessarily have a positive range of

energies and will therefore show the radioactive property. The *normal state* of a non-disintegrated nucleus is then defined as the least energetic of the states of weak quantization which we are about to investigate.

31c. Resonant Energy Intervals.—There are two fundamental procedures for the study of the problem in hand. One of these is to introduce a pair of modified problems having the potential energies indicated by the graphs of $V_1(r)$ and $V_2(r)$ in Fig. 13. Then the V_1 problem will have a set of discrete eigenfunctions located almost entirely within the original crater, while the V_2 problem will have a purely continuous spectrum whose eigenfunctions hardly penetrate the crater at all. The actual problem may be regarded as a combination of the two others with coupling, and the process of radioactive disintegration interpreted as consisting of transitions between almost orthogonal states of the same energy. From this point of view the problem is amenable to perturbation methods similar to those developed in Chap. XII. A great advantage of the method is that it yields a common basis for attack on all problems of weak quantization, including, for example, Oppenheimer's first work on the spontaneous ionization of atoms in electric fields[1] and the theory of the broadening of energy levels by the radiation process.[2]

The second procedure, used by Gamow and by Gurney and Condon, is based on direct application of the B. W. K. approximation method. This procedure lends itself to graphical presentation and is adopted here.

The differential equations to be solved in dealing with our model are of the type already encountered in our initial study of the two-particle problem in Sec. 28. States of any angular momentum which leaves an adequate minimum in the effective potential-energy function are possible. The radial equation can be written in the form

$$H_l \mathfrak{R} = E \mathfrak{R} \tag{31.1}$$

where H_l is the *effective radial Hamiltonian operator* defined by

$$H_l = -\frac{1}{\kappa}\frac{d^2}{dr^2} + \left[V(r) + \frac{l(l+1)}{\kappa r^2} \right] \times. \tag{31.2}$$

Here κ is again the product of $8\pi^2/h^2$ into the reduced mass of the system, but the latter is no longer approximately equal to the electronic mass. We must give it the much larger value $\mu = m_\alpha(M - m_\alpha)/M$, where m_α is the mass of the alpha particle and M is the total mass of the disintegrating nucleus.[3]

---

[1] J. R. OPPENHEIMER, *Phys. Rev.* **31**, 66 (1928). This seems to be the first application of quantum mechanics to a problem of this type.

[2] *Cf.* V. WEISSKOPF and E. WIGNER, *Zeits. f. Physik* **63**, 54 (1930).

[3] The reader should note at this point, if he has not already done so, that the use of the nonrelativistic Equation (31.1) commits us to the approximation of neglecting the variation in the mass of the alpha particle with velocity.

The boundary condition to be used at the origin is the requirement that $\mathfrak{R}(0) = 0$. As no solution of (31·1) for $E > 0$ is quadratically integrable at infinity, it follows from Sec. 30 that all positive energies belong to the continuous spectrum, and that every function $\mathfrak{R}_l(r,E)$, which vanishes at the origin, is a type B eigenfunction. For large values of r these functions reduce to sine waves of length $h/p(\infty,E) = h/(2\mu E)^{1/2}$. The amplitude at infinity of the normalized real waves in which we are chiefly interested is $2^{3/4}(\mu/h^2E)^{1/4}$ (cf. Sec. 30d). Thus it decreases regularly with increasing energy. Inside the crater, however, the variation of amplitude with energy is of a different character. In the latter region there is a resonance phenomenon which is the fundamental cause of the weak quantization.

In order to fix our ideas let us assume that we have to do with a special problem in which the azimuthal quantum number l is zero. The curves $y = E$ and $y = V(r)$ will then cross at two classical turning points r_1 and r_2, shown in Fig. 13, and similar to the turning points x_1 and x_2 of Fig. 7. (When l does not vanish, the effective potential $V + \dfrac{l(l + 1)}{\kappa r^2}$ has a pole of the second order at the origin and there are either three classical turning points, or only one.) Let I, II, III denote respectively the intervals $0 < r < r_1$, $r_1 < r < r_2$, $r_2 < r$. Tracing the course of the integral curve representing a type B eigenfunction, we note that it oscillates in I and III, but is convex to the r axis in II, as if strongly repelled by that axis. On account of the latter property $|\mathfrak{R}(r,E)|$ becomes much larger near the turning point r_2 than in I, except for energies in the neighborhood of those which yield monotonically decreasing values of $|\mathfrak{R}|$ throughout the region II. The special energies in question are evidently approximately the same as the discrete eigenvalues of the modified problem previously mentioned with V replaced by V_1. For these energies the amplitude in the crater is large compared with that outside the crater, and hence large in an absolute sense if the amplitude at infinity is fixed by the normalization condition. For energies much removed from these *resonance* values the amplitude inside the crater is necessarily small.

It is of interest to employ the B. W. K. method to locate the resonance maxima and the widths of the peaks on the resonance curve. For this purpose we assume that the first and second derivatives of V vanish at the origin. Then the discrepancy function Q of (21·12) vanishes at the origin. In the neighborhood of the origin

$$\mathfrak{R}(r,E) = 2B(E)p(r,E)^{-1/2} \cos\left[\frac{2\pi}{h}\int_0^r |p(\zeta,E)|d\zeta - \frac{\pi}{2}\right]$$

$$= 2B(E)p(r,E)^{-1/2} \cos\left[\frac{2\pi}{h}\int_r^{r_1} |p(\zeta,E)|d\zeta + \frac{\pi}{2} - \frac{\pi J(E)}{h}\right], \quad (31·3)$$

where $B(E)$ is a normalizing factor to be determined, and $J(E)$ denotes the Sommerfeld phase integral $2\int_0^{r_1}|p(\zeta,E)|d\zeta$. In order to get a connection between the intervals I and II we shall assume that the "mountain" to be tunneled through is high enough above the energy E so that the B. W. K. approximations are good in the middle of the region II. We now make use of the connection formulas (21·9) and (21·31). The former must be used "against the arrow," which becomes possible if we introduce a small unknown phase angle $\epsilon(E)$ and rewrite (21·9) in the form

$$\underbrace{2p^{-\frac{1}{2}}\cos\left[\frac{2\pi}{h}\int_r^{r_1}|p|d\zeta - \frac{\pi}{4} + \epsilon\right]}_{\text{I}} \leftrightarrow \underbrace{|p|^{-\frac{1}{2}}e^{-\frac{2\pi}{h}\int_{r_1}^r|p|d\zeta}}_{\text{II}}. \qquad (31\cdot4)$$

Let Eq. (31·3) be rewritten in the form

$$\mathfrak{R}(r,E) = 2Bp^{-\frac{1}{2}}\left[\cos\left(\frac{\pi J}{h} - \frac{\pi}{4} + \epsilon\right)\cos\left(\frac{2\pi}{h}\int_r^{r_1}|p|d\zeta + \frac{\pi}{4} + \epsilon\right)\right.$$
$$\left. + \sin\left(\frac{\pi J}{h} - \frac{\pi}{4} + \epsilon\right)\cos\left(\frac{2\pi}{h}\int_r^{r_1}|p|d\zeta - \frac{\pi}{4} + \epsilon\right)\right]. \qquad (31\cdot5)$$

Equations (31·4) and (21·31) now give for the *good* portion of II (*i.e.*, the part of the interval in which $|Q/p^2| \ll 1$)

$$\mathfrak{R}(r,E)_{\text{II}} = B|p|^{-\frac{1}{2}}\left[2\cos\left(\pi\frac{J}{h} - \frac{\pi}{4} + \epsilon\right)e^{\frac{2\pi}{h}\int_{r_1}^r|p|d\zeta}\right.$$
$$\left. + \sin\left(\frac{\pi J}{h} - \frac{\pi}{4} + \epsilon\right)e^{-\frac{2\pi}{h}\int_{r_1}^r|p|d\zeta}\right]. \qquad (31\cdot6)$$

From the above equation it is evident that if $\cos(\pi J/h - \pi/4 + \epsilon)$ is zero, $|\mathfrak{R}(r,E)_{\text{II}}|$ will decrease monotonically throughout the good portion of II. We infer that the amplitude of the oscillations in I will be large compared with those in II for some energy values not very different from those given by the condition that $\pi J/h - \pi/4$ shall be an odd multiple of $\pi/2$. In other words an approximate condition for resonance is that

$$J(E) = (n + \tfrac{3}{4})h. \qquad (31\cdot7)$$

This is just the B. W. K. approximation condition for the location of the discrete eigenvalues of the modified problem with V replaced by V_1. The appearance of the "quarter-integral" quantum number is due to the special boundary condition at the origin and has nothing to do with the problem of weak quantization.

In order to complete our task we have to bridge the gap between II and III with the aid of connection formulas (21·8) and (21·30), or an equivalent. As the latter formula would have to be used "against the

arrow," it is simpler to make direct application of the matrix connection formulas (21·25). Let K denote the integral $\dfrac{2\pi}{h}\displaystyle\int_{r_1}^{r_2}|p|d\zeta$ and let φ stand for the quantity $(\pi J/h - \pi/4 + \epsilon)$. Choosing that branch of the function $p^{-\frac12}$ which makes the product $p^{-\frac12}e^{i\pi/4}$ real in II, we readily reduce the expression for $\Re(r,E)_{\mathrm{II}}$ to the appropriate standard form for the connection formulas, viz.,

$$\Re(r,E)_{\mathrm{II}} = Be^{\frac{i\pi}{4}}\!\left(2\,\cos\,\varphi\,e^{K}p^{-\frac12}e^{-\frac{2\pi}{h}\int_{r}^{r_2}|p|d\zeta} + \sin\,\varphi\,e^{-K}p^{-\frac12}e^{\frac{2\pi}{h}\int_{r}^{r_2}|p|d\zeta}\right).$$

(31·8)

This is of the form $\alpha_u f_u + \alpha_v f_v$ with f_u and f_v defined as in Sec. 21e and the coefficients

$$\alpha_u = B\,\sin\,\varphi\,e^{-K+\frac{i\pi}{4}}; \qquad \alpha_v = 2B\,\cos\,\varphi\,e^{K+\frac{i\pi}{4}}.$$

(31·9)

The corresponding coefficients in III are accordingly determined by

$$\beta_u = \beta_v{}^* = Be^{\frac{i\pi}{4}}[\tfrac12\,\sin\,\varphi\,e^{-K} + i(X\,\sin\,\varphi\,e^{-K} - 2\,\cos\,\varphi\,e^{K})],$$

where iX is the unknown imaginary part of g_{uu} [cf. Eq. (21·29), however]. The radial wave function in III is consequently expressible as

$$\Re(r,E)_{\mathrm{III}} = |\beta_u|p^{-\frac12}\cos\!\left(\frac{2\pi}{h}\int_{r_2}^{r}pd\zeta + \gamma\right).$$

(31·10)

We need not concern ourselves with the phase angle γ. $|\beta_u|$ is to be normalized to the standard value $2^{\frac34}(\mu/h^2E)^{\frac14}$. Hence

$$\frac{1}{B^2} = \left(\frac{h^2E}{8\mu}\right)^{\frac12}\left[\frac{\sin^2\varphi\,e^{-2K}}{4} + (X\,\sin\,\varphi\,e^{-K} - 2\,\cos\,\varphi\,e^{K})^2\right].$$

(31·11)

In spite of the uncertainty in the value of X, this equation enables us to determine the position and character of the resonance regions. $|X|$ is of course less than $|g_{uu}|$ and, if there is a good path Γ through the complex r plane connecting the good portion of II with the good portion of III, it follows from the inequality (21·29) that $|g_{uu}|$ is very much less than e^{K}. Thus $|Xe^{-K}| \ll 1$. Inspection of (31·11) shows that B can be large only when $\cos\,\varphi$, the cofactor of the very large quantity e^{K} on the right, is very nearly zero. Hence we can set $\sin\,\varphi$ equal to $+1$ without appreciable error in the range of resonance values with which we are concerned. Then

$$B^{-2} = \left(\frac{h^2E}{8\mu}\right)^{\frac12}\left[e^{-2K}\!\left(X^2 + \frac14\right) - 4X\,\cos\,\varphi + 4e^{2K}\cos^2\varphi\right].$$

(31·12)

Except in the immediate neighborhood of points for which $\cos\,\varphi = 0$, the last term is by far the largest and its variation with φ will evidently

be the controlling factor in determining the position and shape of the resonance bands. To get the point of maximum resonance we have to set the derivative of the right-hand member of (31·12) with respect to the energy equal to zero.

In calculating this derivative we can assume that X is insensitive to small energy changes. In fact, it would be completely independent of the energy if the potential $V(r)$ were exactly linear over the range of r values bridged by the connection formulas in passing from II to III. Furthermore, the derivatives of $\varphi = \dfrac{\pi J}{h} - \dfrac{\pi}{4} + \epsilon$ and K are of the same order of magnitude. Bearing these items in mind it is possible to neglect a number of terms in the derivative and one readily verifies that, to a close approximation,

$$\frac{d}{dE}\left(\frac{1}{B^2}\right) = \left(\frac{h^2 E}{8\mu}\right)^{1/2}\frac{4\pi}{h}\frac{dJ}{dE}(2\cos\varphi\, e^{2K} - X).$$

Setting this derivative equal to zero, we find that the condition for maximum resonance is that

$$\cos\varphi \equiv \cos\left(\frac{\pi J}{h} - \frac{\pi}{4} + \epsilon\right) = \frac{X}{2}e^{-2K}. \qquad (31·13)$$

As the right-hand member of this equation is small compared with unity when a good path around r_2 in the complex plane exists, we may regard this equation as an approximate confirmation of (31·7). It does not seem possible to improve on (31·7) with the mathematical tools we are using. The value of B^{-2} for the resonance position $E = E_k$ is

$$B^{-2} = \left(\frac{h^2 E_k}{8\mu}\right)^{1/2}\frac{e^{-2K(E_k)}}{4}. \qquad (31·14)$$

In other words B^2 is $4e^{2K}$ times as great as the corresponding squared amplitude at infinity.

We proceed to the computation of the half breadth ΔE_k, of a resonance peak, which we define as the absolute value of the energy change required to reduce B^2 from its maximum value to half maximum value. In making the calculation we shall treat X, K, and the factor $(h^2 E/8\mu)^{1/2}$ on the right side of (31·12) as constants over the resonance interval. Inserting for B^{-2} in (31·12) twice the value given by (31·14), we obtain

$$\frac{e^{-2K}}{4} = (Xe^{-K} - 2\cos\varphi\, e^{K})^2. \qquad (31·15)$$

Let φ_m and $\varphi_{1/2}$ denote the roots of (31·13) and (31·15), respectively, which belong to the resonance region of E_k. Then these equations give

$$\cos\varphi_{1/2} = \cos\varphi_m \pm \frac{e^{-2K}}{4}.$$

The desired half breadth is accordingly

$$\Delta E_k \cong \frac{e^{-2K}}{4} \left(\frac{dE}{d \cos \varphi} \right)_{\cos \varphi = 0} = \frac{he^{-2K}}{4\pi} \left(\frac{dE}{dJ} \right)_{E = E_k}.$$

The quantity dE/dJ which appears in this equation is well known to be the classical frequency $\omega(E)$ of a particle vibrating between the turning points $r = 0$ and $r = r_1$ under the influence of the potential energy $V(r)$. Hence our result can be given the final form

$$\Delta E_k = h\omega(E_k)\frac{e^{-2K(E_k)}}{4\pi}. \tag{31.16}$$

In practice the number $e^{-2K(E_k)}$ is exceedingly small, although it becomes equal to unity for an energy E_k which touches the rim of the crater. The variations in $\omega(E)$ can be neglected in comparison with those of its cofactor in (31·16). Thus Eqs. (31·14) and (31·16) point to exceedingly sharp resonance with high maximum amplitude ratios. It will be observed that the upper resonance peaks are broader than the lower ones.

31d. Energy Distribution in Weakly Quantized States.—It will be immediately evident that the existence of these resonance energies implies the existence of corresponding *approximately monochromatic*, quadratically integrable solutions of the Schrödinger equation $H\Psi = -\dfrac{h}{2\pi i}\cdot\dfrac{\partial \Psi}{\partial t}$, having the property that $|\Psi|^2$ is nearly independent of time. One can set up the wave function of such a quasi-monochromatic quasi-stationary state in various ways, one of which is to choose an initial wave form which is quadratically integrable but nevertheless everywhere an approximate solution of $H\Psi = E\Psi$ for a resonance energy $E = E_k$. It will suffice to set the angular momentum equal to zero, in which case the wave function reduces to a radial factor which we designate as $G(r,t)$ and the operator H_l to H_0 [*cf.* Eq. (31·1)]. In other words, $H\Psi = -\dfrac{h}{2\pi i}\dfrac{\partial \Psi}{\partial t}$ reduces to $H_0 G(r,t) = -\dfrac{h}{2\pi i}\dfrac{\partial G}{\partial t}$.

Let us now identify $G(r,0)$ with the kth discrete eigenfunction of $H_0'u = Eu$, choosing for H_0' the radial Hamiltonian of the modified problem previously mentioned in which the potential function of the model, *viz.*, $V(r)$, has been replaced by a function $V_1(r)$ which is equal to V from the origin to the apex of the crater but continues to rise indefinitely as r increases beyond that point. Such an initial function $u_k(r)$ will be sensibly equal to zero for all values of r for which H_0' differs appreciably from H_0. Hence it is an approximate solution of $H_0\mathfrak{R} = E_k\mathfrak{R}$ for all values of r in the sense that $(H_0 - E_k)u_k$ is everywhere small compared with the amplitude of u_k inside the crater.

Our initial surmise that the probability density $|\Psi|^2$ will be nearly constant in time is supported by a computation of the derivatives of $|G(r,t)|^2$ at the initial instant when G reduces to u_k. For example,

$$\left(\frac{\partial}{\partial t}|G|^2\right)_{t=0} = \frac{2\pi i}{h}(u_k H_0 u_k{}^* - u_k{}^* H_0 u_k)$$

$$= \frac{2\pi i}{h}[u_k(H_0 - E_k)u_k{}^* - u_k{}^*(H_0 - E_k)u_k] \cong 0. \quad (31 \cdot 17)$$

Postponing a fuller discussion of the variation of the wave function with time, we pause to consider the frequency, or energy, distribution which follows from our choice of $G(r,0)$. To this end it is necessary to express $G(r,t)$ in the form

$$G(r,t) = \int_{-\infty}^{+\infty} \alpha(E)\mathfrak{R}(r,E)e^{-\frac{2\pi i E t}{h}} dE, \quad (31 \cdot 18)$$

where $\mathfrak{R}(r,E)$ is a normalized continuous-spectrum eigenfunction of $(31 \cdot 1)$. By $(30 \cdot 25)$,

$$\alpha(E) = \int_0^\infty G(r,0)\mathfrak{R}(r,E)dr = \int_0^\infty u_k \mathfrak{R}(r,E)dr. \quad (31 \cdot 19)$$

$\alpha(E)$ will clearly be negligible except in that particular resonance region to which u_k belongs. For this narrow range of energies and for that range of r values for which u_k is appreciably different from zero, we can treat $\mathfrak{R}(r,E)$ as a multiple of u_k. Furthermore, the ratio of $\mathfrak{R}(r,E)$ to u_k in the intervals under consideration must be proportional to $B(E)$. Thus, to a close approximation,

$$\alpha(E) = gB(E)\int_0^\infty u_k{}^2 dr = gB(E), \quad (31 \cdot 20)$$

where g is a constant with which we are not concerned.

In accordance with optical analogy we may assume that $|\alpha(E)|^2 dE$ is a measure of the probability of the energy interval dE for systems in a state described by $G(r,t)$. This postulate will be more fully justified in the discussions of Secs. 35 and 36. If the postulate be granted, we can infer that the energy distribution in the weakly quantized state is given by $B(E)^2$ to a constant factor of proportionality. Hence we can identify the quantity ΔE_k with the approximate uncertainty of the energy associated with this state.

31e. The Disintegration Process.—The final step in our discussion is to show that, if we start with a weakly quantized state, the amplitude of the wave function inside the crater decays exponentially in time, while the integrated intensity beyond the barrier increases to compensate. When this has been proved we shall have verified the experimental decay law for alpha-particle disintegration.

The problem can be attacked in a number of ways of which that suggested by von Laue[1] is the simplest. Laue makes use of a quasi-classical argument. From the prequantum point of view we should expect from our model that the alpha particle would vibrate back and forth between $r = 0$ and $r = r_1$ with a frequency $\omega(E_k)$ until the time that it actually passes through the barrier. Our study of the transmission of matter waves through potential barriers has shown that in the case of a well-rounded potential hill the transmission coefficient is $T = 1/(1 + e^{2K})$. We should accordingly expect that at each impact the classical particle would have a chance $1/(1 + e^{2K})$ of escape. Consider an assemblage of similar models of which $N(t)$ remain in the initial not disintegrated state at the time t. The number which escape per second will then be equal to $N(t)/(1 + e^{2K})$ multiplied by the number of impacts per second $\omega(E_k)$. Thus we arrive at the rule

$$\frac{dN(t)}{dt} = -\frac{N(t)\omega(E_k)}{[1 + e^{2K(E_k)}]}.$$

Integrating, we obtain

$$N(t) = N(0)e^{-\lambda t}; \qquad \lambda = \frac{\omega(E_k)}{[1 + e^{2K(E_k)}]}. \tag{31.21}$$

This expression for the decay constant λ is perhaps as good as any. The formula makes λ depend primarily on $K(E_k)$. Let us assume that the potential functions $V(r)$ vary rather slowly with the atomic number of the atom under consideration. Then the change of K from one set of alpha particles to another will be due primarily to the change of K with energy rather than with Z. In the case of a parabolic hill—which the actual barrier probably approximates very poorly—dK/dE is constant, being equal to a multiple of the period of vibration of a classical particle moving under the influence of a potential $-V(r)$ and therefore subject to an elastic restoring force. For the values of K actually to be considered $e^{2K} \gg 1$. Hence, in roughest first approximation,

$$\frac{1}{2\lambda}\frac{d\lambda}{dE} = -\frac{dK}{dE} = \text{constant} > 0. \tag{31.22}$$

As the energy E of the initial weakly quantized state is the same as the energy of the escaping alpha particles, the above equation is seen to be an approximate form of the empirical Geiger-Nuttall relation. The theoretical inaccuracy of this relation is paralleled by its experimental inaccuracy.

Another, and theoretically somewhat more satisfactory, method of attack sticks to the wave point of view throughout, although making some use of physical and mathematical intuition. We know from Sec.

---

[1] M. von Laue, *Zeits. f. Physik* **52**, 726 (1928).

21j that in the case of a potential barrier in one dimension flanked by regions of positive kinetic energy there are solutions of the equation $H\psi = E\psi$ for every energy, which can be interpreted as descriptions of steady primary streams of particles incident on the barrier from the left or right, as the case may be, together with corresponding reflected and transmitted streams. The notion of mass or probability current density is equally useful in connection with our nuclear model. Let $\Psi = G(r,t)/r$ denote a normalized solution of the second Schrödinger equation for the problem representing a state of zero angular momentum. Then

$$\frac{d}{dt}\int_0^{r'} r^2 dr \int_0^{\pi} \sin\theta \, d\theta \int_0^{2\pi} |\Psi|^2 d\varphi = 4\pi\frac{d}{dt}\int_0^{r'} |G|^2 dr$$

$$= -\frac{h}{\mu i}\left(G^*\frac{\partial G}{\partial r} - G\frac{\partial G^*}{\partial r}\right)_{r=r'}. \quad (31\cdot23)$$

Let $\mathbb{C}[G]$ denote the function $\frac{h}{\mu i}\left(G^*\frac{\partial G}{\partial r} - G\frac{\partial G^*}{\partial r}\right)$. It follows that in this case we can identify $\mathbb{C}[G]_{r'}$ with the instantaneous total current of probability crossing a sphere of radius r' in the outward direction. Similarly, in the case of a solution of the radial equation $H_0\mathfrak{R} = E\mathfrak{R}$ we can identify $\mathbb{C}[\mathfrak{R}]$ with the relative radial current of an hypothetical infinite stream of particles (*cf.* Sec. 8, p. 31).

Let $f_u(r,E)$ and $f_v(r,E)$ denote the B. W. K. approximations

$$f_u = p^{-\frac{1}{2}}e^{\frac{2\pi i}{h}\int_{r_1}^{r} p(\xi,E)d\xi}; \qquad f_v = p^{\frac{1}{2}}e^{-\frac{2\pi i}{h}\int_{r_1}^{r} p(\xi,E)d\xi},$$

using the same conventions for the evaluation of p and $p^{-\frac{1}{2}}$ from the equation $p^2 = 2\mu(E - V)$ as in Sec. 21j. Then we see from Sec. 21j that a solution of the radial equation which has the form $\mathfrak{R}(r,E) = f_u(r,E)$ near the origin can be interpreted as a stream of particles moving away from the barrier, *i.e.*, *inward*, in that region and having a relative current value $4\pi/\mu$. Similarly an f_v solution represents an equal stream moving *outward* from the origin toward the barrier. In any region where p is real, the current, for a linear combination $a_u f_u + a_v f_v$ with constant coefficients, is equal to the sum of the currents for the two terms taken separately. Hence, we can regard such a function as a description of a superposition of two streams moving in opposite directions. Outside the volcano in III the same interpretation is possible except that f_u gives an outward radial current of magnitude $4\pi e^{-2K}/\mu$ and f_v an inward radial current of magnitude $4\pi e^{2K}/\mu$. In the case of a continuous-spectrum eigenfunction the requirement that $\mathfrak{R}(r,E)$ shall vanish at the origin provides for total reflection at the origin and makes the net current zero. If $\mathfrak{R}(r,E)$ has a larger amplitude in I than in III, the diverging and converging partial currents inside the crater will be correspondingly

large compared with the partial currents outside. Whether this condition will obtain, or its reverse, is evidently a matter of phases. If we set E equal to the eigenvalue E_k of the V_1 problem, the phases are such as approximately to minimize the external currents relative to the internal currents.

Consider next the wave function $G(r,t)$ generated by a weakly quantized initial state $G(r,0)$. In Sec. 31d we found it convenient to identify $G(r,0)$ with the kth discrete eigenfunction of the modified V_1 problem whose equation is $H_0'u = Eu$. For our present purpose it is better to give $G(r,0)$ some such form as

$$G(r,0) = e^{-\sigma r^n}\mathfrak{R}(r,E_k),$$

choosing n as a positive integer equal to or greater than 2, and giving σ a value such that the modulating factor $e^{-\sigma r^n}$ differs but little from unity at the turning point $r = r_1$ yet approaches zero rather rapidly outside the crater. $\mathfrak{R}(r,E_k)$ is here a real class B eigenfunction of the radial equation. If the barrier is a high one, we can fix n and σ so that $G(r,0)$ is sensibly equal to $\mathfrak{R}(r,E_k)$ inside the crater, whereas the chance that the particle is initially outside the crater is very small. The definition of $G(r,0)$ will then differ so little from that previously used that the analysis of Sec. 31d will be unaffected.

Physical—or is it mathematical?—intuition now tells us what to expect for the behavior of $G(r,t)$ in time with such an initial state; and some readers, at any rate, will take a greater satisfaction in finding an intuitional explanation of the observed phenomenon than in finding a formal mathematical proof without an intuitional background.

At the initial instant we have inside the crater a superposition of two sets of approximately monochromatic waves of equal amplitude progressing in opposite directions.[1] There is total reflection at the origin and very nearly total, though gradual, reflection by the crater wall. The small outward transmission through this wall is compensated by a reverse transmission from similar progressive waves on the outside. The latter, however, are damped radially and are essentially finite wave trains. We may accordingly expect the front of the outward-moving train (the f_u train) to progress away from the origin with a speed equal to the group velocity while the tail of the inward-moving train contracts toward the origin with the same speed. Thus the incoming transmission through the barrier will quickly disappear and we shall have a continual unbalanced outward flow of probability through the crater wall. If the ratio of the net outward current computed for, say, the outer turning point r_2, to the integrated wave intensity inside the crater were constant,

---

[1] The progress is to be measured by the group velocity rather than by the phase velocity.

the probability that the alpha particle is inside the crater would decay exponentially in time. In that case we should have

$$-\frac{d}{dt} \log \int_0^{r_2} |G|^2 dr = \frac{\mathcal{C}[G]_{r_2}}{4\pi \int_0^{r_2} |G|^2 dr} = \lambda = \text{constant.} \qquad (31\cdot24)$$

Actually this ratio will not be exactly constant, but it will be very nearly so, for by Sec. 31d, and for points inside the crater, we can replace $\mathfrak{R}(r,E)$ to a close approximation by $\mathfrak{R}(r,E_k)$ in (31·18). $G(r,t)$ immediately breaks into the product of $\mathfrak{R}(r,E_k)$ and a function of the time, say $\varphi(t)$. Thus the wave form of the waves trapped inside the crater must be nearly constant in time, and the transmission coefficient for the outgoing waves incident on the barrier must be at all times very nearly equal to the transmission coefficient for a train of waves of uniform amplitude and energy E_k incident on that barrier. The emergent current is then equal to the current reflected from the origin multiplied by $1/(1 + e^{2K})$. To get the latter current we must resolve $G(r,t)$ in the neighborhood of the origin into the linear combination

$$G(r,t) = a_u f_u + a_v f_v. \qquad (31\cdot25)$$

The outward current at the origin is then $\mathcal{C}[a_u f_u] = \dfrac{4\pi}{\mu} |a_u|^2$. Thus

$$\lambda = \frac{|a_u|^2}{(1 + e^{2K})} \left[\mu \int_0^{r_2} |G|^2 dr \right]^{-1}.$$

To complete the computation $\int_0^{r_2} |G|^2 dr$ must be evaluated in terms of $|a_u|^2$. For this purpose it is convenient to reduce (31·25) to the form

$$G = e^{ix} |a_u| |p^{-\frac{1}{2}}| \sin\left[\frac{2\pi}{h} \int_0^r |p| d\zeta \right]. \qquad (31\cdot26)$$

This equation is the initial condition for the exact function $G(r,t)$, which is an approximate multiple of $\mathfrak{R}(r,E_k)$, and so an approximate solution of $H_0 \mathfrak{R} = E_k \mathfrak{R}$. An exact evaluation of G and the desired integral is unnecessary since the law of force is known only in a qualitative manner, and the model itself is only a crude one. For the *upper* nuclear-energy levels E_k the approximation of Sec. 21i is appropriate and yields

$$\int_0^{r_2} |G|^2 dr \cong \int_0^{r_1} |G|^2 dr = \frac{|a_u|^2}{2} \int_0^{r} \frac{dr}{p} = \frac{|a_u|^2}{\mu\omega(E_k)}, \qquad (31\cdot27)$$

where $\omega(E_k)$ is the classical radial-vibration frequency inside the crater. Thus we obtain the value $\omega(E_k)/(1 + e^{2K})$ for λ in exact agreement with the von Laue result (31·21).

In the case of the normal state of the nucleus, in which we are chiefly interested, the foregoing approximation is unsatisfactory although correct as regards order of magnitude. For this lowest energy level E_0 the function G will have no nodes between the origin and r_2. Since it starts off at the origin as a sine curve with amplitude $|a_u||p(0,E_0)^{-\frac{1}{2}}|$, it is reasonable to evaluate the integral, treating G as a single arch of a sine curve with the above amplitude and a phase $3\pi/4$ at $r = r_1$. On this hypothesis $\int |G|^2 dr = 2r_1|a_u|^2/3p(0,E_0)$ and λ has the value $\dfrac{3p(0,E_0)/2r_1\mu}{(1 + e^{2K})}$.

It is of interest to note that the von Laue value of the decay constant when combined with the estimated energy uncertainty of a weakly quantized state gives an interesting illustration of the Heisenberg uncertainty principle. In a time $1/\lambda$ the probability that an initially undisintegrated nucleus will have disintegrated is reduced to $1/e$. Hence we can say that the uncertainty in the lifetime of an undisintegrated state is of the order of magnitude of $1/\lambda$. As the product of $1/\lambda$ and ΔE_k, using (31·21) and (31·16), is $h/4\pi$ with an error of the order of the small quantity e^{-2K}, we see that we have to do with an almost ideal case in which the uncertainty in energy and time has a minimum product.

*31f. Complex Eigenvalues.—On account of its mathematical interest, mention should be made here of the elegant method of complex eigenvalues applied to the decay problem by Gamow.[1] The method starts from the observation that the experimental law of radioactive decay would be *exactly* described by a wave function of the form

$$\Psi = \frac{\Re(r)e^{-\frac{\lambda t}{2} + \frac{2\pi i E t}{h}}}{r} \tag{31·28}$$

A function of this type will satisfy the second Schrödinger equation, if $\Re(r)$ is a solution of the radial equation

$$H_0\Re = (E - i\alpha)\Re. \qquad \lambda = \frac{4\pi\alpha}{h} \tag{31·29}$$

This latter equation with the suggested complex value of the energy parameter is of course a perfectly good differential equation and has the same sort of manifold of solutions as the corresponding equation for a real parameter. The solutions can be studied by the B. W. K. method as before.

In this case the function $p(x,\ E - i\alpha) = \sqrt{2\mu(E - i\alpha - V)}$ is complex on both sides of the potential barrier as well as in the region II "under the mountain." However, if we take into account the experimental values of the decay constant λ and the energy of the escaping alpha particles, we must assume that in all practical cases the imaginary

[1] *Loc. cit.*, footnote 2, p. 179.

part of p is extremely small close to the zeros of its real part. Consider the case of thorium emanation, the atoms of which have an average life of the order of 1 min. λ has the value 1.27×10^{-2} sec.$^{-1}$ while the energy is 10^{-5} ergs. Hence $\alpha/E = h\lambda/4\pi E = 6.6 \times 10^{-24}$. We can accordingly neglect α^2 in comparison with $(E - V)^2$, thus obtaining the approximation

$$p = p_0 - \frac{\iota\mu\alpha}{p_0}; \qquad p_0 = [2\mu(E - V)]^{\frac{1}{2}}. \qquad (31\cdot30)$$

We identify r_1 and r_2 in this case with the zeros of p_0. In the region III the B. W. K. approximations f_u and f_v become

$$f_u = p^{-\frac{1}{2}} \exp\left\{\frac{2\pi i}{h}\int_{r_1}^{r_2} p\, d\zeta + \frac{2\pi i}{h}\int_{r_2}^{r} p_0\, d\zeta + \frac{2\pi\mu\alpha}{h}\int_{r_2}^{r}\frac{d\zeta}{p_0}\right\},$$

$$f_v = p^{-\frac{1}{2}} \exp\left\{-\frac{2\pi i}{h}\int_{r_1}^{r_2} p\, d\zeta - \frac{2\pi i}{h}\int_{r_2}^{r} p_0\, d\zeta - \frac{2\pi\mu\alpha}{h}\int_{r_2}^{r}\frac{d\zeta}{p_0}\right\}. \qquad (31\cdot31)$$

As p_0 is real and positive in III by our conventions, $|f_u|$ increases without limit as r increases, while $|f_v|$ decreases and approaches zero at $r = \infty$. The current theorem of Eq. (31·23) is still valid and the currents of the f_u and f_v approximations are sensibly the same as before, except for an additional multiplying factor due to the contribution of α to the variation in the amplitudes with r. The important point for us to note is that f_u still represents an outward flow of current and f_v an inward one.

Any solution of (31·29) yields a corresponding function Ψ given by Eq. (31·28) which can be interpreted as the superposition of two decaying streams of alpha particles, one radiating from the origin and the other converging toward it. These currents can be balanced at the origin if the absolute value of the f_u component of $\mathfrak{R}$ is equal to the absolute value of the f_v component at the same point. This condition can be satisfied for any value of the complex number $E - i\alpha$. In general, however, a solution $\mathfrak{R}(r, E - i\alpha)$ will yield converging as well as diverging currents in III and thus will differ radically from the form needed for the representation of an assemblage of disintegrating nuclei. There exist solutions of (31·29), however, which give a net current zero at the origin and also take the asymptotic form $a_u f_u(r, E - i\alpha)$ in III. A necessary condition that $\mathfrak{R}$ shall be of this type is that the current reflected by the crater shall differ from the diverging current incident on the crater by an amount equal to the transmitted current. But by using a small positive value of α we provide for a slight increase in the diverging current between the origin and the crater and also a slight decrease in the converging current between the same pair of points. For just one value of α the current difference at the crater will just balance the leakage due to the tunnel effect. This balancing of currents is insufficient, however, to insure that there

shall be no converging stream outside the hill. For that purpose it is necessary that $\mathcal{R}$ shall have a definite phase at the tunnel entrance r_1. When this phase is properly chosen the phase at the origin will not in general make the origin a nodal point for $\mathcal{R}$. In order to keep Ψ from becoming infinite at the origin we must then choose the real energy E in a special way so that the phase requirement at the crater will be compatible with the requirement that $\mathcal{R}$ vanish at the origin. Owing to the very small transmission coefficient, α is very small, and hence the energies E_k' which conform to the above phase requirement are very nearly the same as the energies E_k of our preceding discussion. The successive values of $E_k' - i\alpha_k'$ for the different resonance points on the energy scale are called *complex eigenvalues* of the energy, although they have not at all the same significance as real discrete eigenvalues.

In the neighborhood of the origin the eigenfunctions $\mathcal{R}(r, E_k' - i\alpha_k')$ conform to the requirements for the initial weakly quantized state of an assemblage of predisintegrated nuclei. Moreover, the corresponding complete wave function $\Psi(r,t)$ decays in time according to the exponential law found by experiment. Hence Gamow assumes the legitimacy of identifying the decay constant of an observed nuclear disintegration process with the decay constant λ_k computed by means of (31·29) from a corresponding complex eigenvalue $E_k' - i\alpha_k'$.

Let us grant for the moment the legitimacy of the Gamow hypothesis. We can then identify E_k' with E_k with sufficient precision to permit an accurate estimate of α_k'. Equation (31·24) is valid if we identify G with $\mathcal{R}(r, E_k' - i\alpha_k')e^{-\frac{2\pi i}{h}(E_k' - i\alpha_k')t}$. Hence

$$\lambda = \frac{h}{4\pi\mu i} \frac{\left[\mathcal{R}^*\dfrac{\partial\mathcal{R}}{\partial r} - \mathcal{R}\dfrac{\partial\mathcal{R}^*}{\partial r}\right]_{r=r_1}}{\displaystyle\int_0^{r_1}|\mathcal{R}|^2\,dr}. \tag{31·32}$$

In evaluating λ by this formula we ought, strictly speaking, to use the eigenfunction $\mathcal{R}(r, E_k' - i\alpha_k')$, but on account of the extreme smallness of α_k' it is evident that there will be no appreciable error if we substitute for this eigenfunction the function $\mathcal{R}(r,E_k)$ used in Sec. 31e. Thus we come back in first approximation to the result previously derived from an altogether different point of view. Higher approximations could be obtained but are not worth while.

In view of the agreement with the result obtained in Sec. 31e and that derived by the perturbation-theory method the legitimacy of the Gamow hypothesis can hardly be questioned. Unfortunately, however, a completely satisfactory justification of the method independent of any alternative method of approach is lacking. To be sure we can create a

quadratically integrable wave function representing a weakly quantized initial state by multiplying $\Re(r, E_k' - i\alpha_k')$ by a modulating factor, say $e^{-\sigma r^n}$, and can then justify (31·32) by the same type of heuristic argument as we have used in Sec. 31e. The problem of placing these heuristic arguments on a rigorous basis has not been solved as yet, however.

In concluding this section we call the reader's attention to two other important examples of weak quantization.[1] The so-called *predissociated molecular states* are imperfectly quantized excited states of molecules having energies greater than the minimum dissociation energy of the molecule and having a finite lifetime due to spontaneous transitions to the dissociated states. *The energy levels which give rise to X-ray line spectra* are the energies of weakly quantized ionic states capable of a second spontaneous ionization in which the energy set free by an outer electron dropping down into a vacant space in an "inner shell" is used to eject a second outer electron. This type of process, called the *Auger effect*, will be the subject of further discussion in Sec. 32h.

Radiative transitions can take place between different imperfectly quantized states and between imperfectly quantized states and sharply quantized states. The diffuseness of the initial or final energy level, as the case may be, then produces a broadening of the spectrum line emitted or absorbed and referable ultimately to the finite lifetime of one or both of the associated energy levels. A measurable broadening of this type is found in many band-spectrum lines having predissociated initial or final states and in many X-ray lines, especially those of longer wave length.

In a sense all atomic and molecular energy levels except normal states and metastable states—and the latter form a doubtful exception—are imperfectly quantized, for the existence of spontaneous radiative transitions from an upper energy level to any lower level implies a finite lifetime for the former. The broadening of spectrum lines produced in this way by the emission of radiation itself is swamped in most cases, however, by broadening due to other causes such as collisions, Doppler effect, etc.

32. THE EXISTENCE AND PROPERTIES OF SOLUTIONS OF THE MANY-PARTICLE SCHRÖDINGER EIGENVALUE-EIGENFUNCTION PROBLEM

32a. Introduction.—In the work of Secs. 28 to 30 we have used the method of the separation of variables to demonstrate the existence and properties of solutions of the Schrödinger eigenvalue-eigenfunction problem for two oppositely charged interacting particles. The separation of variables in effect resolves the three-dimensional two-particle problem

---

[1] *Cf.* O. K. RICE, *Phys. Rev.* **34**, 1451 (1929), **35**, 1538 (1930) for a theoretical discussion of predissociated molecules with references to the literature. For theory of the Auger effect *cf.*, *e.g.*, G. WENTZEL, *Zeits. f. Physik* **43**, 524 (1927), *Physik. Zeits.* **29**, 333 (1928); E. FUES, *Zeits. f. Physik* **43**, 726 (1927).

into three Sturm-Liouville problems and permits us to use the highly developed Sturm-Liouville theory in this more general case. The wave equation for two particles with like charges can be separated in the same way, although the radial equation has no discrete eigenvalues. The method is inapplicable, on the other hand, to the more complicated problems of atomic and molecular structure involving three or more interacting particles, and, in fact, the existence of solutions of these problems has yet to be proved.

The basic differential equation for the problem of $n = f + 1$ particles is (17·1). Let $\overrightarrow{r_k}$ denote the radius vector from particle $f + 1$ to particle k, and let x_k, y_k, z_k denote its components. Let r_{jk} denote the absolute value of the distance from particle j to particle k. Assuming that all particles are subject to interactions of the Coulomb or electrostatic type,[1] we throw Eq. (17·1) into the form

$$
H_r\psi \equiv \left\{ -\frac{h^2}{8\pi^2}\left[\sum_{k=1}^{f}\frac{1}{\mu_k}\left(\frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2}\right) \right.\right.
$$
$$
\left. + \frac{1}{\mu_{f+1}}\sum_{k=1}^{f}\sum_{j=1}^{f}\left(\frac{\partial^2}{\partial x_k \partial x_j} + \frac{\partial^2}{\partial y_k \partial y_j} + \frac{\partial^2}{\partial z_k \partial z_j}\right) \right]
$$
$$
\left. + \sum_{k=1}^{f}\frac{e_k e_{f+1}}{r_k} + \sum_{k=1}^{f-1}\sum_{j=k+1}^{f}\frac{e_k e_j}{r_{jk}}\right\}\psi = E\psi. \quad (32\cdot1)
$$

If this equation is correct, it should have quadratically integrable eigenfunctions corresponding to the stable states of atoms, molecules, and ions revealed by experiment. Moreover, an essentially arbitrary function of the coordinates representing the instantaneous form of a wave packet should be capable of expansion into a discrete-continuous linear combination of type A and type B eigenfunctions. In Secs. 19 and 23 we have developed existence theorems for quadratically integrable eigenfunctions in one dimension, but no corresponding existence theorem for the many-dimensional case has been constructed. In fact when we pass from the two-particle problem to the many-particle problem of (32·1) we pass from a domain in which there is a well-developed basic mathematical theory to a domain of mathematical ignorance. It becomes necessary to *assume* both the existence and basic properties of the discrete- and continuous-spectrum eigenfunctions. These assumptions can be made out of hand, or they can be partially justified by plausibility considerations of a mathematical character. We adopt the

---

[1] An assumption which must be discarded when a satisfactory relativistic formation of the quantum theory is finally constructed.

second procedure, which at least exhibits the difficulties in the nature of a satisfactory general theory if it does not overcome them.

It will be of comfort to the reader of the remainder of this section to note that we have some evidence of the existence of discrete eigenfunctions for the simpler many-dimensional problems through the success of attempts to work them out and locate the eigenvalues by successive approximations. We have no proof of the ultimate convergence of these approximations, but in the best work of this kind on the helium atom[1] and the hydrogen molecule,[2] for example, the computed approximate energies appear to converge in a quite satisfactory manner upon values very close to the experimental ones.

32b. New Boundary Conditions for Physically Admissible Wave Functions.—Before attempting a formal extension of the general theory of one-dimensional eigenvalue problems to many dimensions we pause to reconsider the boundary-continuity conditions for physically admissible wave functions. In laying down our preliminary definition of physically admissible wave functions (type A functions) in Sec. 17 we were primarily concerned with the selection of the discrete eigenfunctions of $H\psi = E\psi$ from the totality of the solutions of this equation. We should like, however, to define a class of "physically admissible" functions which shall include the above mentioned eigenfunctions and which are to be admitted as descriptions of the most general subjective states, or of corresponding assemblages of identical systems prepared so as to be in a common state. We shall have much to do with the business of expanding such functions in terms of the eigenfunctions of various operators, and it will reduce our worries if we can assume that all allowed ψ functions satisfy rather stringent boundary-continuity conditions.

As the solutions of the hydrogenic-atom problem worked out in Sec. 29 are particularly well behaved both at the origin and at infinity, the question arises whether we cannot require that all physically admissible wave functions shall share these desirable features. Such a restriction is evidently permissible provided that the discrete eigenfunctions of the many-particle problem have the same characteristics as those of the two-particle problem, and provided that the proposed restrictions on physically admissible wave functions leave a class broad enough to describe any experimentally realizable physical state.

Clearly the best way to answer our question is to begin by defining the proposed restrictions, reserving for later discussion the question of their validity. We accordingly lay down a series of five conditions D_1, D_2, D_3, D_4, D_5 and designate a function which conforms to all of them as a function of class or type D. The phrase "physically admissible" will be interpreted for the present as synonymous with "type D." In

[1] E. A. HYLLERAAS, *Zeits. f. Physik* **54**, 347 (1929).

[2] H. L. JAMES and A. S. COOLIDGE, *J. Chem. Phys.* **1**, 825 (1933).

Sec. 42b we shall add a further restriction to the manifold of physically admissible wave functions not included in the definition of D. The conditions are defined with reference to a particular Hamiltonian operator, so that each Hamiltonian generates its own class D, and are directly applicable in general only when the functions are expressed in terms of Cartesian coordinates. They are also limited to Hamiltonians of the usual nonrelativistic type with no other singularities than those produced by simple Coulomb type poles in the potential function [*cf.* Eq. (32·1)].[1]

The explicit formulation of the D conditions is as follows.

D_1.—Every $\psi(x_1, \cdots, z_f)$ in class D is single-valued and analytic in all the variables at every point where the potential energy $V(x_1, \cdots, z_f)$ is analytic. In other words ψ is analytic at every finite point of configuration space that does not bring two charged particles together.

D_2.—ψ shall vanish at infinity faster than any negative power of the coordinates. This requirement which is very simple in the one-dimensional case can be formulated more precisely as follows: Let Q denote a point in the coordinate space and let P be any polynomial of the coordinates $x_1, \cdots, z_f$. Then, if Q moves out to infinity along any path involving no singular points,

$$\lim_{Q \to \infty} P(Q)\psi(Q) = 0.$$

This condition is assumed to apply as well to the first and second derivatives of ψ with respect to the coordinates of the configuration space.

D_3.—ψ and its first derivatives shall be absolutely and quadratically integrable over the whole configuration space.[2]

D_4.—In order to insure that the probability current to the singular domain $r_{ij} = 0$ shall vanish, it is necessary to introduce a condition best

---

[1] This limitation in applicability is necessary because the discrete eigenfunctions of an equation of the form $H\psi = E\psi$ always have "built in" singularities fitted to the singularities in H. These are of an entirely different character in the relativistic and nonrelativistic theories of the hydrogen atom.

In the case of a negative inverse-square potential the radial equation of the two-particle problem has no proper set of eigenfunctions at all. In this case none of the solutions of the radial equation satisfies the singular-point boundary conditions. The general problem of the inverse-square potential has been discussed by Shortley, *Phys. Rev.* **38**, 120 (1931) but without reference to the singular-point boundary conditions. The difficulties to which it leads are of no real physical importance since the inverse-square potential does not exist in nature. In fact we must regard even the Hamiltonian of (32·1) with its Coulomb potential terms as a useful and convenient approximation to the true Hamiltonian operator rather than anything of absolute significance.

[2] This evidently implies that the product of $\psi(x_1, \cdots, z_f)$ or $\dfrac{\partial \psi}{\partial x_j}$ $(j = 1, 2, \cdots)$ with any polynomial $P_n(x_1, \cdots, z_f)$ is also absolutely and quadratically integrable since these polynomials are analytic and bounded in every finite domain. The rapid decrease of ψ and its derivatives take care of the convergence at infinity.

described in terms of a new set of coordinates. In place of the coordinates x_i, y_i, z_i, x_j, y_j, z_j, we introduce the coordinates ξ, η, ζ of the center of gravity of particles i and j together with the spherical relative coordinates r_{ij}, θ_{ij}, φ_{ij}. Let $\Phi(r_{ij}, \theta_{ij}, \varphi_{ij})$ denote the integral

$$\Phi = \int_{\infty} \cdots \int |\psi|^2 d\xi d\eta d\zeta d\tau',$$

in which $d\tau'$ is the product of the differentials of all the Cartesian coordinates except x_i, y_i, z_i, x_j, y_j, z_j. The quadratic integrability of ψ insures the convergence of the integral. Let A denote the sphere $r_{ij} = a$ in the three-dimensional space of r_{ij}, θ_{ij}, φ_{ij}, and let W_a denote the mean value of Φ for the sphere A. Thus

$$W_a[\psi] \equiv \frac{1}{4\pi a^2} \int \int_A \Phi a^2 \sin \theta_{ij} \, d\theta_{ij} d\varphi_{ij} \equiv \frac{1}{4\pi} \int \int_A \Phi d\Omega.$$

Similarly we designate by $W_a \left[\dfrac{\partial \psi}{\partial r_{ij}} \right]$ the corresponding mean with $\partial \psi / \partial r_{ij}$ substituted for ψ throughout.

We now require that if ψ is in class D there shall exist a positive real number ϵ and a real number m such that $a^{2-m} W_a[\psi]$ and $a^{2+m-\epsilon} W_a \left[\dfrac{\partial \psi}{\partial r_{ij}} \right]$ are bounded in the neighborhood of $a = 0$.

D_5.—If ψ is in D, the conditions D_1 to D_4 shall apply to $H^n \psi$ as well as to ψ itself. Here H^n denotes the n-fold application of the Hamiltonian operator H. The series $\displaystyle\sum_{n=0}^{\infty} \left(-\frac{2\pi it}{h} \right)^n \frac{H^n \psi}{n!}$ is assumed to converge for all values of t in the neighborhood of every point $x_1, \cdots, z_f$ and must yield a function which is also of class D.

The reader will readily verify that these conditions are satisfied by the discrete eigenfunctions of the hydrogenic-atom problem and by their linear combinations. Hence there is an infinite multiplicity of class D solutions of the Schrödinger equation $H\Psi = -\dfrac{h}{2\pi i} \dfrac{\partial \Psi}{\partial t}$ for this problem. The existence of class D functions for the Hamiltonian of a many-particle problem has not been proved but is a plausible assumption to be considered below.

We proceed to a brief discussion of the successive items which constitute the definition of class D.

D_1 is an extension of the continuity condition for type A and type B functions. It seems probable that all bounded solutions of the Schrödinger equation $H\Psi = -\dfrac{h}{2\pi i} \dfrac{\partial \Psi}{\partial t}$ are analytic in t and in the space coordi-

nates except on singular domains of the operator H. At any rate this is true of all the familiar solutions of this equation. We wish every "physically admissible" function to be a potential initial state for a physically admissible solution of the above equation and accordingly introduce the condition D_1. The condition is convenient, and, in the writer's opinion, unobjectionable, whether necessary or not.

As regards the adequacy of analytic functions for the description of experimental situations, it may be observed that, although the idealization of physical experiments may lead to situations requiring nonanalytic functions for their description, this is evidently not true for the actual experiments themselves. For example, the exact location of an electron in space could be described in the language of wave mechanics only by a function—the Dirac δ function—so discontinuous that it does not exist! But when we consider the inevitable experimental error we see that the result of an actual positional observation is to give us knowledge about the electron representable by an analytic function which conforms roughly to a condition of the form

$$|\Psi|^2 = Ae^{-\alpha[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]}.$$

In general the interaction of an atomic system with a piece of apparatus conceived of as classical may yield exact information representable only by nonanalytic waves. However, when our imperfect knowledge of the state of the actual instruments is taken into account, we see that the need of nonanalytic functions is apparent rather than real.

The condition D_2 goes beyond the mere requirement of quadratic integrability for type A functions. Together with D_3 it imposes the severest restrictions on the behavior of type D functions at infinity, roughly equivalent to saying that these functions must approach zero exponentially for large values of the coordinates. The requirement of absolute integrability is one which we have already seen to be very convenient in connection with the determination of expansion coefficients where a continuous spectrum is involved (*cf.* pp. 36 and 171).

D_4 is analogous to the s.p.b.c. of Sec. 23*d* and is formulated for the same purpose. Its use permits us to identify Class D with the class of permissible comparison functions for a variational formulation of the problem of locating the discrete eigenvalues of H with their eigenfunctions.

D_5 is designed to insure that the second Schrödinger equation $H\Psi = -\dfrac{h}{2\pi i}\dfrac{\partial \Psi}{\partial t}$ shall transform physically admissible wave functions into new physically admissible wave functions through the passage of time. As previously stated, it seems probable that all bounded solutions of this equation are analytic in the time like the product form functions $\Psi = \psi(x,E)e^{-\frac{2\pi i}{h}Et}$ and their linear combinations. We assume that such

is the case in order to validate our proof (*cf.* footnote 2, p. 18) that Ψ is uniquely determined for all time by the Schrödinger equation and its form at some arbitrary initial instant. It follows from the equation itself that

$$\left[\frac{\partial^n \Psi}{\partial t^n}\right]_{t=0} = -\left(\frac{2\pi i}{h}\right)^n H^n[\Psi]_{t=0}.$$

Thus the series specified in D_5 is the Taylor's series expansion of Ψ in powers of t which must converge for small values of t if Ψ is analytic in t as well as in the space variables.

Let us define the operators $e^{\tau\frac{\partial}{\partial t}}$, $e^{-\frac{2\pi i}{h}\tau H}$ by means of the formal power-series expansion for e^x. If $\Psi(x,t)$ is analytic in t

$$\Psi(x, t + \tau) = \sum_{n=0}^{\infty} \frac{\tau^n}{n!} \frac{\partial^n}{\partial t^n} \Psi(x,t) = e^{\tau\frac{\partial}{\partial t}} \Psi(x,t).$$

If $\Psi(x,t)$ is a solution of the second Schrödinger equation this becomes[1]

$$\Psi(x, t + \tau) = e^{-\frac{2\pi i}{h}\tau H} \Psi(x,t).$$

Hence $e^{-\frac{2\pi i}{h}tH}$ is sometimes called the *time-displacement operator*. By applying this operator to the initial function $\psi(x)$ we obtain a corresponding solution of the second Schrödinger equation. D_5 requires that this operator shall always transform class D functions into class D functions.

An important and readily verified general property of the class D functions is that they form a *linear manifold*. That is, if any two functions ψ_1, ψ_2 conform to the D conditions, an arbitrary linear combination will also conform to these conditions. This is the principle of the "addition of states" emphasized by Dirac.[2]

*32c. Approximating Arbitrary Quadratically Integrable Functions by Means of Class D Functions.—For the special case of three-dimensional functions where class D is defined with reference to the Hamiltonian

$$H = -\frac{h^2}{8\pi^2}\nabla^2 - \frac{Ze^2}{r}, \qquad (32\cdot2)$$

it is possible to prove another important general property of that class, *viz: For every quadratically integrable function $f(x,y,z)$, whether it belongs to class D, or not, and for every positive constant ϵ, however small, there exists a class D function $\psi(x,y,z)$ such that $\int_{\infty} |f - \psi|^2 dx dy dz < \epsilon$.* In the language of von Neumann, the manifold D is "everywhere dense"

[1] *Cf.* VON NEUMANN, *M.G.Q.*, p. 108.

[2] Dirac, *P.Q.M.*, section 7.

in the Hilbert space of all quadratically integrable functions of x, y, and z.[1]

To prove this proposition we make use of the eigenfunctions of the equation

$$\nabla^2 \psi + \left(E + \frac{1}{r} - r^2\right)\psi = 0. \tag{32·3}$$

This is simply a central-force-field problem with a potential function of the Coulomb type at the origin, but becoming infinite at infinity. Solving by the method of the separation of variables, we get a three-dimensional array of eigenfunctions

$$\psi_{nlm} = R_{nl}(r)\Theta_{lm}(\theta)\Phi_m(\varphi),$$

which are of class D with respect to the Hamiltonians of Eqs. (32·2) and (32·3). Each of the three one-dimensional equations into which the original three-dimensional equation is resolved has a complete set of discrete eigenfunctions. It follows[2] that the product functions form a complete set in the sense that

$$\lim_{\lambda,\mu,\nu \to \infty} \int \left| f - \sum_{m=-\lambda}^{+\lambda} \sum_{l=|m|}^{|m|+\mu} \sum_{n=l+1}^{l+\nu} c_{nlm}\psi_{nlm} \right|^2 r^2 \sin\theta\, dr d\theta d\varphi = 0,$$

$$c_{nlm} = (f, \psi_{nlm}).$$

for every quadratically integrable function $f(x,y,z)$. Then for every f and every ϵ we can choose λ, μ, ν large enough so that, if

$$\psi = \sum_{m=-\lambda}^{+\lambda} \sum_{l=|m|}^{|m|+\mu} \sum_{n=l+1}^{l+\nu} \psi_{nlm}(f, \psi_{nlm})$$

the inequality $\int_\infty |f - \psi|^2 dx dy dz < \epsilon$ will hold.

In the general case of a many-particle problem with one of the basic Hamiltonians of Eqs. (7·2), (7·3), and (32·1), it is not possible to force a separation of variables without dropping the singular domains $r_{kj} = 0$. Hence the above type of proof fails, but there can be little doubt of the validity of the theorem in this general case and we assume it as a postulate. This proposition is the basis of our claim that class D is broad enough for the purposes of quantum mechanics.

32d. Hermitian Character of the Hamiltonian Operator.—It is desirable at this stage in the development of the theory to replace the definitions of the adjoint to a one-dimensional second-order differential operator, and of the self-adjoint property, by new definitions of broader scope applicable to linear operators which are not necessarily of

---

[1] *Cf.* VON NEUMANN, *M.G.Q.*, p. 23.

[2] *Cf.* COURANT-HILBERT, *M.M.P.*, Kap. II, §1, 6.

differential character. These terms will be used hereafter only in the new sense.

Definition:[1] *Let C denote a linear manifold of functions defined and quadratically integrable over a domain M of coordinate space. Let O and $O^\dagger$ denote two linear operators which yield quadratically integrable transforms $O\psi$ and $O^\dagger\psi$ when applied to any function ψ of the manifold C. If*

$$(\psi_1, O^\dagger\psi_2) = (O\psi_1, \psi_2) \tag{32·4}$$

for every pair of functions ψ_1, ψ_2 which belong to C when the domain of integration is extended over M, the operator $O^\dagger$ is said to be adjoint to O with respect to the manifold C and the domain M.

It follows as a corollary that if $O^\dagger$ is adjoint to O, the latter operator must be adjoint to $O^\dagger$.

Definition: If the operator $O^\dagger$ is adjoint to itself (self-adjoint) with respect to a manifold C and a domain M, it is said to be Hermitian with respect to C and M. The relation

$$(\psi_1, O\psi_2) = (O\psi_1, \psi_2) \tag{32·5}$$

must then hold for every pair of functions ψ_1, ψ_2 in C when the integration is extended over M.

Corollary: If the linear operators α and β are both Hermitian with respect to a manifold C and a domain M, $\alpha - i\beta$ is adjoint to $\alpha + i\beta$ with respect to the manifold C and the domain M.

Corollary: A Sturm-Liouville operator

$$\Lambda \equiv \frac{d}{dx}\left[p(x)\frac{d}{dx}\right] - q(x) \times$$

is Hermitian with reference to an interval $a \leqslant x \leqslant b$ in which Λ has no singular points and to any class of functions which are twice differentiable

---

[1] This and the following definition may be compared with the corresponding definitions as laid down by von Neumann, *M.G.Q.*, pp. 48, 50. von Neumann makes no mention of the manifold C and domain M in defining either the adjoint operator $O^\dagger$ or the Hermitian property. This is because he includes in the definition of each operator a linear manifold of functions (or of elements of abstract Hilbert space) on which it may operate, and for which its transforms are required to be quadratically integrable over a fundamental domain M (or to be new elements of abstract Hilbert space). We have defined an operator as merely a rule for transforming one function into another. It follows that what we call a single operator can give rise to a multiplicity of operators in the von Neumann sense by a multiplicity of choices of the manifold of functions on which they are allowed to operate.

It will be observed that in the case of a second-order differential operator in one dimension, if $\Lambda^\dagger$ is adjoint to Λ in the sense of the definition given in footnote 1, p. 122, it is also adjoint to Λ in the sense of the new definition with respect to the domain $a < x < b$ and any linear manifold of functions quadratically integrable in that interval and so defined as to make $G(a) = G(b)$.

in the interval and conform to homogeneous boundary conditions at a and b.

Corollary: A Sturm-Liouville operator Λ is Hermitian with reference to an interval $a < x < b$ bounded by singular points, but having no such points in its interior, and to the class of functions which are twice differentiable in the interval and conform to the singular-point boundary condition at a and at b.

Theorem: The basic Hamiltonian operator of the many-particle problem, viz.,

$$H = -\sum_{k=1}^{n} \frac{h^2}{8\pi^2\mu_k}\left(\frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2}\right) + \sum_{k=1}^{n-1}\sum_{j=k+1}^{n} \frac{e_k e_j}{r_{kj}}, \quad (32\cdot6)$$

is Hermitian with reference to the domain consisting of all coordinate space and to the functions of class D.[1]

To prove the theorem we first introduce the $3n$-dimensional vector $\overrightarrow{B}$ with components $u_1, v_1, w_1, u_2, v_2, \cdots w_n$ along the axes $x_1, y_1, z_1, x_2, \cdots z_n$, where

$$u_k = \frac{h^2}{8\pi^2\mu_k}\left(\psi_1\frac{\partial\psi_2^*}{\partial x_k} - \psi_2^*\frac{\partial\psi_1}{\partial x_k}\right),$$

$$v_k = \frac{h^2}{8\pi^2\mu_k}\left(\psi_1\frac{\partial\psi_2^*}{\partial y_k} - \psi_2^*\frac{\partial\psi_1}{\partial y_k}\right),$$

$$w_k = \frac{h^2}{8\pi^2\mu_k}\left(\psi_1\frac{\partial\psi_2^*}{\partial z_k} - \psi_2^*\frac{\partial\psi_1}{\partial z_k}\right).$$

Then

$$\psi_2^*H\psi_1 - \psi_1 H\psi_2^* = \sum_{k=1}^{f}\left(\frac{\partial u_k}{\partial x_k} + \frac{\partial v_k}{\partial y_k} + \frac{\partial w_k}{\partial z_k}\right). \quad (32\cdot7)$$

Let us next form the integral

$$J_G = \int \cdots \int_G [\psi_2^*H\psi_1 - \psi_1 H\psi_2^*]dx_1 dy_1 \cdots dz_n$$

over a finite region G of coordinate space so chosen that it includes no singular points. Employing the method of Gauss we transform J_G into the surface integral of the normal component of $\overrightarrow{B}$ along the outward normal to the $(f-1)$-dimensional hypersurface of G. Let G be bounded by a set of planes perpendicular to the coordinate axes and by hypercylinders

$$r_{kj}^2 \equiv (x_k - x_j)^2 + (y_k - y_j)^2 + (z_k - z_j)^2 = a^2$$

to shut out the singular domains for which $r_{kj} = 0$.

If the boundary planes are now moved out to infinity at the same time that the radii of the hypercylinders approach zero, the integral J_G approaches

$$(H\psi_1, \psi_2) - (\psi_1, H\psi_2)$$

[1] It is not difficult to extend this theorem to the Hamiltonian of Eq. (7·8) with suitable restrictions on the vector potential $\overrightarrow{a}$. The essential features of the required integral transformation for the three-dimensional case are indicated on p. 31.

as a limit. If we can show that under these circumstances the surface integral approaches zero, provided that ψ_1 and ψ_2 belong to class D, we shall thereby prove our theorem.

It suffices to prove that the integral of the absolute value of the normal component of $\overrightarrow{B}$ over the whole of each of these surfaces approaches zero in the limit. Consider first the contribution of the typical cylinder $r_{kj} = a$. The normal component of $\overrightarrow{B}$ is

$$
\begin{aligned}
B_n &= -\left[u_k \frac{\partial r_{kj}}{\partial x_k} + v_k \frac{\partial r_{kj}}{\partial y_k} + w_k \frac{\partial r_{kj}}{\partial z_k} + u_j \frac{\partial r_{kj}}{\partial x_j} + v_j \frac{\partial r_{kj}}{\partial y_j} + w_j \frac{\partial r_{kj}}{\partial z_j} \right] \\
&= \frac{h^2}{8\pi^2} \frac{\mu_k + \mu_j}{\mu_k \mu_j} \left[\psi_2{}^* \frac{\partial \psi_1}{\partial r_{kj}} - \psi_1 \frac{\partial \psi_2{}^*}{\partial r_{kj}} \right].
\end{aligned}
\tag{32.8}
$$

The right-hand member of (32.8) assumes that ψ_1 and ψ_2 are expressed in terms of the coordinates ξ, η, ζ, r_{kj}, θ_{kj}, φ_{kj} introduced in connection with the boundary condition $\bullet D_4$ (p. 199).

The surface integral of B_n over the hypercylinder S_{kj} for which $r_{kj} = a$ is equal to the difference of the integrals

$$
I_{kj}(a) = \frac{h^2}{8\pi^2} \frac{\mu_k + \mu_j}{\mu_k \mu_j} a^2 \int \cdots \int_{S_{kj}} \psi_2{}^* \frac{\partial \psi_1}{\partial r_{kj}} \sin \theta_{kj} \, d\theta_{kj} d\varphi_{kj} d\xi d\eta d\zeta d\tau',
$$

$$
J_{kj}(a) = \frac{h^2}{8\pi^2} \frac{\mu_k + \mu_j}{\mu_k \mu_j} a^2 \int \cdots \int_{S_{kj}} \psi_1 \frac{\partial \psi_2{}^*}{\partial r_{kj}} \sin \theta_{kj} \, d\theta_{kj} d\varphi_{kj} d\xi d\eta d\zeta d\tau'.
$$

From the inequality of Schwarz, Eq. (22.20), it follows that

$$
\left| \frac{8\pi^2}{h^2} \frac{\mu_k \mu_j}{\mu_k + \mu_j} I_{kj}(a) \right|^2 \leqslant a^4 W_a[\psi_2] W_a\left[\frac{\partial \psi_1}{\partial r_{kj}} \right].
$$

In view of the condition D_4 we conclude from the above inequality that $\lim\limits_{a \to 0} I_{kj} = 0$ if ψ_1 and ψ_2 belong to D. As the same argument applies to J_{kj}, the contribution of S_{kj}—and hence of each of the hypercylinders—to the surface integral approaches zero as its radius approaches zero.

Consider next the integral of the absolute value of the normal component of $\overrightarrow{B}$ over the surface of the box made up of the plane boundaries of G. The integral is the product of the mean value of B_n into the area. The latter is proportional to the $(3n - 1)$th power of the linear dimension of the box and it follows from D_3 that if we expand the box in all directions at the same rate the value of B_n will approach zero more rapidly than any finite inverse power of the linear dimension. Hence the integral over the plane boundary surfaces of G also approaches zero as G is expanded to include all coordinate space. It follows that H is Hermitian with respect to coordinate space and the functions of class D. It is, in fact, Hermitian with respect to a much wider class than D, for we have not used all the class D conditions in this proof.

In Sec. 15 we separated the variables of the Schrödinger equation by introducing relative coordinates in place of the "absolute" coordinates of Eqs. (7.2) and (7.3). The transformation there employed resolves the complete Hamiltonian operator H of Eq. (32.6) into the sum of two operators H_r and H_σ giving the energy of the relative motion (internal energy) and the energy of the motion of the center of mass, respectively. The operator

$$H_0 = -\frac{h^2}{8\pi^2 M}\left(\frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}\right) \tag{32·9}$$

is obviously Hermitian with respect to class D functions in absolute coordinate space and also to class D functions in the three-dimensional space X,Y,Z. It follows at once that the internal-energy operator H_r of Eq. (32·1) is Hermitian with respect to class D functions in absolute coordinate space. It is also Hermitian with respect to class D functions in the space of the relative coordinates since we have only to multiply each function of this type by a class D function of X, Y, Z in order to obtain a class D function in absolute coordinate space.

Theorem: The scalar product (ψ_1,ψ_2) *of any two class D solutions of the second Schrödinger equation*

$$H\Psi = -\frac{h}{2\pi i}\frac{\partial\Psi}{\partial t} \tag{32·10}$$

is constant in time. The theorem is a direct consequence of the Hermitian character of H. Thus

$$\frac{d}{dt}(\Psi_1,\Psi_2) = \int_\infty \cdots \int\left[\Psi_1\frac{\partial\Psi_2{}^*}{\partial t} + \Psi_2{}^*\frac{\partial\Psi_1}{\partial t}\right]d\tau = \frac{2\pi i}{h}[(\Psi_1,H\Psi_2) - (H\Psi_1,\Psi_2)] = 0. \tag{32·11}$$

Corollary: The normalization integral (Ψ,Ψ) for any class D solution of Eq. (32·10) is constant in time, as required by Sec. 8.

Theorem: The eigenvalues of Eq. (32·1) for class D eigenfunctions are all real (cf. Sec. 23). As in the one-dimensional case we have

$$(\psi,H_r\psi) - (H_r\psi,\psi) = (E^* - E)(\psi,\psi) = 0.$$

Since (ψ,ψ) cannot vanish unless ψ is identically zero, it follows that $E^* = E$.

Theorem: Any two class D eigenfunctions of Eq. (32·1), say ψ_1, ψ_2, having different energies E_1, E_2 must be mutually orthogonal. Thus

$$(\psi_1,H_r\psi_2) - (H_r\psi_1,\psi_2) = (E_2 - E_1)(\psi_1,\psi_2) = 0.$$

32e. Reduction of the Eigenvalue-eigenfunction Problem for Discrete Spectra to Variational Form.—Consider the solutions of the variational equation

$$\delta([H_r - E]\psi,\psi) = \delta\int_\infty \psi^*[H_r - E]\psi d\tau = 0, \tag{32·12}$$

with comparison functions subject to the boundary-continuity conditions D.

$$\delta([H_r - E]\psi,\psi) = ([H_r - E]\delta\psi,\psi) + ([H_r - E]\psi,\delta\psi).$$

Since all comparison functions are of class D, ψ and $\delta\psi$ must be of class D.

But H_r is Hermitian with respect to class D. Hence we obtain the reduction

$$\delta([H_r - E]\psi,\psi) = \text{real part } 2\int_\infty \delta\psi^*[H_r - E]\psi d\tau = 0. \quad (32\cdot13)$$

In view of the theorem of Sec. 32c the functions $\delta\psi$ are sufficiently arbitrary so that (32·13) can hold only when ψ is a class D solution of

$$H_r\psi = E\psi.$$

The extremals of the variational problem are the eigenfunctions of the differential equation and *vice versa*.

The above formulation of the variational problem parallels the formulation A of Sec. 24. It is easy to give other formulations parallel to the schemes B and C of Sec. 24. Thus the eigenvalues and eigenfunctions of Eq. (32·1) are the stationary values of Q/N and the extremals which yield those values, respectively, if Q and N are defined by

$$Q \equiv (H_r\psi,\psi); \qquad N \equiv (\psi,\psi). \qquad (32\cdot14)$$

The integral Q when applied to class D functions is readily reduced to the alternative form

$$Q[\psi] \equiv (H_r\psi,\psi) = \int_\infty \left\{ \frac{h^2}{8\pi^2}\left[\sum_{k=1}^{f}\frac{1}{\mu_k}|\text{grad}_k\,\psi|^2 + \frac{1}{\mu_{f+1}}\left(\sum_{k=1}^{f}\frac{\partial\psi^*}{\partial x_k}\right)\left(\sum_{j=1}^{f}\frac{\partial\psi}{\partial x_j}\right) \right.\right.$$
$$\left.\left. + \frac{1}{\mu_{f+1}}\left(\sum_{k=1}^{f}\frac{\partial\psi^*}{\partial y_k}\right)\left(\sum_{j=1}^{f}\frac{\partial\psi}{\partial y_j}\right) + \frac{1}{\mu_{f+1}}\left(\sum_{k=1}^{f}\frac{\partial\psi^*}{\partial z_k}\right)\left(\sum_{j=1}^{f}\frac{\partial\psi}{\partial z_j}\right) \right] \right.$$
$$\left. + \sum_{k=1}^{f}\frac{e_k e_{f+1}}{r_k}\psi\psi^* + \sum_{k\neq j}\frac{e_k e_j}{r_{kj}}\psi\psi^* \right\}d\tau. \quad (32\cdot15)$$

32f. A Lower Bound for the Energy Integral.—The contribution of the kinetic-energy terms in the integrand, *i.e.*, terms in the derivatives of ψ, is essentially positive as is the contribution of the mutual potential energy of pairs of particles carrying charges of like sign. Hence it is possible to show that in the case of an atom or ion with a single positively charged nucleus Q has a lower bound. We identify the nucleus with the particle $f + 1$ and introduce the integral $\tilde{Q}$ defined by

$$\tilde{Q}[\psi] = \int_\infty \left[\frac{h^2}{8\pi^2}\sum_{k=1}^{f}\frac{1}{\mu_k}|\text{grad}_k\,\psi|^2 + \sum_{k=1}^{f}\frac{e_k e_{f+1}}{r_k}\psi\psi^* \right]d\tau. \quad (32\cdot16)$$

$\tilde{Q}$ is obtained from Q by dropping positive terms in the integrand of (32·15). Hence

$$Q[\psi] > \tilde{Q}[\psi]$$

for all functions ψ. It follows that $Q[\psi]/N[\psi]$ has a lower bound if such a bound exists for $\tilde{Q}[\psi]/N[\psi]$. But $\tilde{Q}[\psi]$ is the energy integral for a modified Schrödinger problem involving f *independent* particles moving around a fixed center of force. The variables in this modified problem separate in appropriate coordinates yielding eigenfunctions $\psi_{kn}{}^{(0)}$, $\psi_{kE}{}^{(0)}$ which are products of hydrogenic wave functions. An arbitrary class D function of the coordinates $x_1, \cdots, z_f$ can be expanded in terms of these eigenfunctions. If the expansion is given the form

$$\psi = \sum_k \sum_n c_{kn}\psi_{kn}{}^{(0)} + \sum_k \int c_{kE}\psi_{kE}{}^{(0)}dE,$$

it follows from the completeness of the system of hydrogenic functions that

$$\frac{Q[\psi]}{N[\psi]} \geqslant \frac{\tilde{Q}[\psi]}{N[\psi]} = \frac{\displaystyle\sum_k \left[\sum_n |c_{kn}|^2 E_n{}^{(0)} + \int |c_{kE}|^2 E\,dE \right]}{\displaystyle\sum_k \left[\sum_n |c_{kn}|^2 + \int |c_{kE}|^2 dE \right]} \geqslant E_0{}^{(0)}.$$

Thus the minimum eigenvalue of the modified problem is also a lower bound for Q/N.

The corresponding theorem for the general case of a system containing any number of positively and negatively charged particles does not go through so easily, but we shall assume that it is true.

The conclusion that Q/N has a lower bound makes it reasonable to suppose that there is a function $\psi_0(x_1, \cdots, z_f)$ which actually minimizes Q/N but does not prove the existence of such a function. The existence of a lower bound for Q/N is an obvious *necessary* condition for the existence of the minimizing function but is not *sufficient*. This is proved by the consideration of the problem of the motion of a free particle in one dimension. If the potential energy is set equal to zero, the energy integral is essentially positive, but, although it is bounded below, it has no discrete eigenvalues and no minimizing function.

*32g. Behavior of Solutions of the Differential Equation at Singular Domains.—Another necessary condition for the existence of a minimizing function ψ_0 is that *the differential equation shall have solutions which conform to the class D boundary condition in the neighborhood of each of its singular domains*. The absence of a minimizing function in the above mentioned one-dimensional case may be ascribed to the failure of this second necessary condition.

We proceed to a preliminary examination of solutions of the differential equation in the neighborhood of the singular domains for the special case of the helium atom. The system consists of two electrons and a

nucleus, which we designate as particles 1, 2, 3, respectively. The potential energy is

$$V = -\frac{2e^2}{r_{13}} - \frac{2e^2}{r_{23}} + \frac{e^2}{r_{12}}.$$

The singular domains to be investigated are (1) the domains $r_{13} = 0$, $r_{23} = 0$, $r_{12} = 0$, representing two-particle collisions; (2) the point $r_{13} = r_{23} = r_{12} = 0$, representing a three-particle collision; and (3) the domain at infinity. Different coordinate systems are useful for studying the different singular domains. First, using the scheme mentioned in footnote 1, p. 64 we write the differential equation in the form

$$\left[\frac{\partial^2}{\partial \xi_1^2} + \frac{\partial^2}{\partial \eta_1^2} + \frac{\partial^2}{\partial \zeta_1^2} + \frac{\partial^2}{\partial \xi_2^2} + \frac{\partial^2}{\partial \eta_2^2} + \frac{\partial^2}{\partial \zeta_2^2} \right]\psi + \kappa(E - V)\psi = 0, \quad (32 \cdot 17)$$

with

$$\xi_1 = x_1 - x_3; \qquad \xi_2 = \sqrt{\frac{\sigma}{\mu}}\left(x_2 - \frac{x_1\mu_1 + x_3\mu_3}{\mu_1 + \mu_3} \right);$$

$$\eta_1 = y_1 - y_3; \qquad \eta_2 = \sqrt{\frac{\sigma}{\mu}}\left(y_2 - \frac{y_1\mu_1 + y_3\mu_3}{\mu_1 + \mu_3} \right);$$

$$\zeta_1 = z_1 - z_3; \qquad \zeta_2 = \sqrt{\frac{\sigma}{\mu}}\left(z_2 - \frac{z_1\mu_1 + z_3\mu_3}{\mu_1 + \mu_3} \right);$$

$$\sigma = \frac{\mu_2(\mu_1 + \mu_2)}{\mu_1 + \mu_2 + \mu_3}; \qquad \mu = \frac{\mu_1\mu_3}{\mu_1 + \mu_3}; \qquad \kappa = \frac{8\pi^2\mu}{h^2}.$$

If we make the approximation of treating the nuclear mass μ_3 as infinite in comparison with the electronic mass, ξ_2, η_2, ζ_2 reduce to the x, y, z components of r_{23}, respectively. Introducing spherical coordinates ρ_1, θ_1, φ_1 in place of ξ_1, η_1, ζ_1, we reduce the equation to the form

$$\frac{1}{\rho_1^2} \frac{\partial}{\partial \rho_1}\left(\rho_1^2 \frac{\partial \psi}{\partial \rho_1} \right) + \frac{1}{\rho_1^2} \Lambda_1\psi + \nabla_2^2\psi + \kappa(E - V)\psi = 0,$$

where Λ_1 and ∇_2^2 are the operators

$$\Lambda_1 = \frac{1}{\sin \theta_1} \frac{\partial}{\partial \theta_1}\left(\sin \theta_1 \frac{\partial}{\partial \theta_1} \right) + \frac{1}{\sin^2\theta_1} \frac{\partial}{\partial \varphi_1},$$

$$\nabla_2^2 = \frac{\partial^2}{\partial \xi_2^2} + \frac{\partial^2}{\partial \eta_2^2} + \frac{\partial^2}{\partial \zeta_2^2}.$$

If we now enter the differential equation with the expansion

$$\psi = \sum_{n=0}^{\infty} \rho_1^n F_n(\theta_1, \varphi_1, \xi_2, \eta_2, \zeta_2),$$

we obtain a sequence of partial differential equations for the coefficients $F_n(\theta_1, \varphi_1, \xi_2, \eta_2, \zeta_2)$. There is no apparent difficulty about solving these

equations so long as we restrict the domain of the expansion to the region in which r_{13}, *i.e.*, ρ_1, is less than r_{23}. Thus we conclude that (32·17) has solutions which are finite on the singular domains (1).

If the interelectronic energy term e^2/r_{12} is neglected, the variables can be separated in Eq. (32·17), and solutions are obtainable in the form of the product of two three-dimensional hydrogenic-atom eigenfunctions. These product eigenfunctions for the modified problem can be written as power series in the six-dimensional radius

$$R = [\xi_1{}^2 + \eta_1{}^2 + \zeta_1{}^2 + \xi_2{}^2 + \eta_2{}^2 + \zeta_2{}^2]^{\frac{1}{2}}.$$

Hence we are led to investigate the behavior of solutions of the original problem near the singular point (2) by seeking solutions which are power series in R. Introducing six-dimensional spherical coordinates, defined by

$$\xi_1 = R \cos \alpha \sin \theta_1 \cos \varphi_1, \qquad \xi_2 = R \sin \alpha \sin \theta_2 \cos \varphi_2,$$
$$\eta_1 = R \cos \alpha \sin \theta_1 \sin \varphi_1, \qquad \eta_2 = R \sin \alpha \sin \theta_2 \sin \varphi_2,$$
$$\zeta_1 = R \cos \alpha \cos \theta_1, \qquad \zeta_2 = R \sin \alpha \cos \theta,$$

we convert (32·17) into the form

$$\frac{1}{R^5} \frac{\partial}{\partial R}\left(R^5 \frac{\partial \psi}{\partial R}\right) + \frac{1}{R^2}D\psi + \kappa(E - V)\psi = 0, \qquad (32\cdot18)$$

where D is a Hermitian linear operator involving the angles α, θ_1, θ_2, φ_1, φ_2 and independent of R. Its explicit form is given in Eq. (35·14). The potential function V is of the form U/R, where U depends only on $\alpha, \theta_1, \cdots, \varphi_2$.

If a solution exists of the form

$$\psi = R^\beta \sum_{n=0}^{\infty} R^n G_n(\alpha,\theta_1,\varphi_1,\theta_2,\varphi_2), \qquad (32\cdot19)$$

it is necessary that the functions G_n shall belong to the Hermitian manifold of D and that in consequence β shall have the value zero. Unfortunately this substitution yields a sequence of equations for the successive G_n's which have not been solved and in all probability have no solutions in the desired Hermitian manifold. Thus one can say with considerable confidence that if eigenfunctions of (32·17) exist, the singular point at the origin lies outside the scope of (32·19). We have no answer, however, to the question whether, or not, there are solutions of (32·17) having a more complicated behavior, but still falling within the Hermitian manifold of the Hamiltonian.[1]

The disappointment of the above conclusion is somewhat relieved when we reflect that the Coulomb potential function used in our Hamil-

---

[1] The problem is under investigation by Prof. J. H. Bartlett to whom the author is indebted for valuable suggestions. See two articles on the helium-wave equation by T. H. Gronwall and J. H. Bartlett, *Phys. Rev.* **51**, 655, 661 (1937).

tonian is not known *a priori* to be absolutely correct. It would be possible to invent a substitute potential function analytic over all coordinate space and yet so nearly equal to the Coulomb function, except in the immediate neighborhood of the domains (1) and (2), as to be experimentally indistinguishable from the latter. Eigenfunctions of a Schrödinger equation using such a modified potential function certainly do exist. Therefore an ultimate negative conclusion regarding our existence theorem would mean a slight modification of the Coulomb law rather than a fundamental change in the theory of wave mechanics. For the present it is convenient to assume what we cannot prove, *viz.*, the existence of solutions of (32·17) which do conform to the boundary condition D_4 and are consequently adapted to the construction of eigenfunctions.

A similar attack can be made on the singular points $r_1 = 0$, $r_2 = 0$, $r_{12} = 0$, using coordinates appropriate to power-series expansions in r_1, r_2, or r_{12} as the case may be, and with similar semisatisfactory results.

Consider next the behavior of solutions of the differential equation (32·18) for very large values of R. Except in those singular directions along which one of the three terms in the potential energy is infinite, we can neglect $V\psi$. The variables are then separable and the hypothesis that $\psi = u(R)v(\alpha_1 \cdots \varphi_2)$ gives for every negative energy an infinity of approximate solutions of (32·17) which approach zero at infinity as $e^{-\sqrt{-\kappa E}R}$ and which therefore conform to the D conditions at infinity. Solutions of the form $\psi = uv$ for positive values of E are not quadratically integrable. We infer that the helium atom can have no discrete positive-energy eigenvalues, but that, so far as this portion of the boundary is concerned, any negative energy gives solutions of the differential equation which conform to the class D boundary conditions.

Thus we find essentially the same difference in the behavior of solutions near infinity for positive and negative energies as in the three-dimensional hydrogenic atom. One might readily jump at the conclusion that here again there is a discrete spectrum of negative-energy eigenvalues which meets a continuous spectrum of positive energies at $E = 0$. This conclusion must be wrong, however, for experimentally the continuous spectrum extends down to the energy of the normal state of the He+ ion, *i.e.*, to $-4Rh$. An examination of the behavior of solutions for large values of r_2 and r_{12} throws some light on the problem.

In this domain we can neglect the potential-energy terms in $1/r_2$ and $1/r_{12}$, treating electron 2 as a free particle. The Schrödinger equation in the coordinates of (32·17) reduces to

$$(\nabla_1{}^2 + \nabla_2{}^2)\psi + \kappa\left(E + \frac{2e^2}{r_1}\right)\psi = 0. \tag{32·20}$$

It has solutions of the form $\psi = \psi_1(\xi_1,\eta_1,\zeta_1)\psi_2(\xi_2,\eta_2,\zeta_2)$, where ψ_1 and ψ_2 are solutions of the equations

$$\nabla_1^2\psi_1 + \kappa\left(E_1 + \frac{2e^2}{r_1}\right)\psi_1 = 0, \tag{32.21}$$

$$\nabla_2^2\psi_2 + \kappa(E - E_1)\psi_2 = 0. \tag{32.22}$$

The first of this pair of equations is that of the He$^+$ ion, while the second is of the free-particle type.

In that portion of configuration space in which e^2/r_2 and e^2/r_{12} are much smaller than E, any solution of the initial Eq. (32.17), say $\psi(\xi_1, \cdots ,\zeta_2,E)$ must reduce to a linear combination of products of solutions ψ_1, ψ_2 of the above equations. If ψ is of class D, or merely quadratically integrable, each term of this linear combination must be quadratically integrable over that portion of configuration space in which the neglected terms of the potential energy are actually small. But this means that ψ_1 is quadratically integrable over *all* ξ_1, η_1, ζ_1 space. In fact, if ψ is of class D, ψ_1 must be a class D eigenfunction of (32.21) and E_1 a discrete (negative) energy level of the He$^+$ ion. ψ_2 in turn must be quadratically integrable over all that portion of ξ_2, η_2, ζ_2 space in which r_2 is large, a condition which can be met by giving $E - E_1$ any negative value—as may be seen by introducing spherical coordinates. Thus it seems possible to find functions $\psi(\xi_1, \cdots ,\zeta_2,E)$ for every negative value of E which, in the region where e^2/r_2 and e^2/r_{12} are small, conform to all boundary conditions appropriate to a discrete eigenfunction except those which apply at $r_{12} = 0$ and $r_2 = 0$.

Consider next, however, the possibility of a type B eigenfunction (*i.e.*, one which satisfies the D conditions at finite points but is not quadratically integrable) of (32.17) with a negative energy E. If such an eigenfunction exists, it must also be expressible as a linear combination of products $\psi_1\psi_2$ in each of which ψ_1 must be either a class D or class B eigenfunction of (32.21) and ψ_2 must be bounded in the neighborhood of infinity. In the former case E_1 is a helium-ion energy level as before, but E_2 must be positive. In the latter case E_1 is positive, and if we make the approximation of treating the nuclear mass as infinite in comparison with the electronic mass, we can show that E_2 must be a He$^+$ energy level.[1] There is in either case the possibility of class B eigenfunctions for energies as low as that of the normal state of He$^+$ which we designate as $E_N{}^+$. We are thus confronted with an apparent overlapping of the discrete and continuous spectra in the energy interval $E_N{}^+ < E < 0$.

---

[1] Neglecting e^2/r_{12}, but keeping both of the other terms in the potential energy, one obtains a He$^+$ equation for each of the two electrons. The approximations are valid over enough of coordinate space to prove that if the energy of either electron is negative it must be a discrete energy level.

*32h. The Auger Effect.—Such an overlapping *can* take place for a properly chosen Hamiltonian and does certainly occur in the case of a pair of electrons moving around a common center of force but having no interactions with each other. Moreover, we know by experiment that, in general, atoms and atomic ions do have relatively sharp emission and absorption lines associated with energy levels which lie above the lower limit of the continuous energy-level spectrum, *i.e.*, above the lowest ionization potential. These overlapping energy levels are apparently of the weakly quantized variety, however, atoms in such states passing spontaneously without emission or absorption of radiation into ionized (continuous-spectrum) states of the same energy. This process, known as the *Auger effect*,[1] is entirely analogous to the spontaneous disintegration of atomic nuclei by the emission of alpha particles (*cf.* Sec. 31). It was discovered in the study of Wilson cloud chamber photographs of gases irradiated by X-rays. In addition to the long tracks of photo-electrons, shorter tracks were found which could be accounted for only as the result of spontaneous transitions from excited singly ionized states (due to the removal of inner electrons) to states of equal energy consisting of a doubly ionized atom and an ejected electron. More recently it has been shown[2] that the great breadth of many X-ray lines is probably due in large part to the shortening of the lifetime of the associated energy levels by the Auger effect. There is also definite evidence of similar spontaneous ionization from the upper energy levels of the optical spectra of certain atoms. In the light of Sec. 31 we must interpret these results as proof that, strictly speaking, the discrete and continuous spectra of atoms do not ordinarily overlap, the energy levels from which spontaneous ionization takes place being associated with states which are imperfectly quantized. In other words, we associate the Auger effect with quadratically integrable functions which for certain fairly well defined energies yield approximate solutions of the first Schrödinger equation and hence give rise to quasi-stationary states of finite, but relatively long, lifetime, capable of emitting and absorbing radiation like true discrete energy-level states. It may happen, of course, that the effect of the emission and absorption of radiation on the lifetime of an imperfectly quantized state which is neglected in setting up the basic equation (32·1) is greater than the effect of spontaneous ionization. In that case a distinction between exactly quantized and imperfectly quantized states based on (32·1) becomes somewhat academic.

Returning from these general considerations to the helium-atom problem, we note that our examination of the domain in which r_2 and r_{12} are large is insufficient to prove the existence of *any* discrete energy

[1] P. Auger, *Ann. de Physique* **6**, 183 (1926).
[2] E. Ramberg and F. K. Richtmyer, *Phys. Rev.* **47**, 644, 805 (1935).

levels, to say nothing of giving a definite proof of the existence of exact class D eigenfunctions, for energies above the lowest ionization potential. It does show, however, that there are two possible ways in which a bounded solution of the wave equation with an energy in the interval $E_N^+ < E < 0$ can behave in the region where r_2 and r_{12} are large, leaving it an open question whether the boundary conditions for r_2 and r_{12} small can be met independently by wave functions having these two types of behavior, or whether wave functions which satisfy the above mentioned "inner" boundary conditions must always reduce at infinity to linear combinations of the two types. Theoretical considerations can be urged against the former alternative but are not sufficiently conclusive to warrant reproduction here. Granting the validity of the second alternative on the basis of experiment, we may properly reinterpret the original argument for an overlapping of the discrete and continuous spectra as theoretical evidence of the existence of weakly quantized states in the energy region $E_N^+ < E < 0$.

32i. The Discrete Eigenfunctions of the Differential Equation as Minimizing Functions.—In Sec. 32e we saw that the general atomic eigenvalue-eigenfunction problem can be reduced to variational form; in Sec. 32f we proved that the quantity $Q[\psi]/N[\psi]$ to be varied has a lower bound $E_0^{(0)}$; and in Secs. 32g and 32h we satisfied ourselves, for the special case of the helium problem, that solutions of the Schrödinger equation which conform to the class D boundary conditions exist in the neighborhood of each singular domain provided only that the energy is less than the energy of the normal state of the ionized atom, E_N^+. If it can be proved that E_N^+ is greater than $E_0^{(0)}$, we shall know that all the necessary conditions for the existence of a function of class D, which actually minimizes Q/N, are satisfied. Actually $E_0^{(0)}$ and E_N^+ are readily worked out in this case, each being derivable from the solution of a hydrogenic-atom problem. We find that $E_N^+ = -4Rh$, and $E_0^{(0)} = -8Rh < E_N^+$, where R is the Rydberg constant.

At this point we cross the Rubicon by assuming, not only for helium, but for every stable atomic or molecular system of particles, that there exists a class D function ψ_0 which minimizes Q/N and gives it a value E_0 which we identify with the energy of the normal state of the system. We further assume that in general there exists a sequence of class D functions $\psi_0, \psi_1, \psi_2, \cdots$ having the property that each minimizes Q/N subject to the class D boundary-continuity conditions and to the additional restriction that all admissible comparison functions shall be orthogonal to each earlier function of the sequence. Since each set of comparison functions is more restricted than the set used in the preceding problem, the eigenvalues $E_0, E_1, \cdots, E_n, \cdots$ conform to the inequality $E_n \geqslant E_{n-1}$. If two or more of the eigenvalues are equal, we have a degenerate energy level. (The numbering of the eigenvalues, it will

be observed, corresponds to that of an ordered sequence of minimizing functions and *not* to the number of distinct energy values.) By Sec. 32e the eigenvalues and eigenfunctions obtained in this way are the eigenvalues and eigenfunctions of the appropriate Schrödinger equation, (32·1). From a theorem on p. 206 we know that eigenfunctions belonging to *different* discrete eigenvalues must always be orthogonal. However, *when they are derived in this way, independent eigenfunctions having the same eigenvalue are also orthogonal.* On the other hand, since any arbitrary linear combination of a set of degenerate eigenfunctions of the Schrödinger equation is an eigenfunction of the same eigenvalue, the orthogonality property is not a necessary characteristic of an arbitrary set of degenerate eigenfunctions.

32j. The Continuous Spectrum and the Completeness of the System of Eigenfunctions.—If the series of discrete eigenvalues were infinite in every case, and if we could assume $\lim_{n \to \infty} E_n = \infty$, the argument of Sec. 25 could be used to prove the completeness of the sequence $\psi_0, \psi_1, \cdots$, $\psi_n \cdots$ and our right to expand an arbitrary class D function into an infinite series of its members. However, as the discrete eigenvalues are negative and, in fact, have been assumed to lie below E_N^+, the class D eigenfunctions do not form a complete set. In order to set up an expansion theorem it is necessary to make use of solutions of the Schrödinger equation which conform to the class D conditions at finite points but which are not quadratically integrable in the neighborhood of infinity. We assume the existence of such type B eigenfunctions for all energies above the energy of the normal state of the once-ionized system. We infer from Sec. 32g that in portions of coordinate space, where one particle has a negligible mutual potential energy with the rest of the system, each of these type B functions can be factored into the product of a positive-energy type-B eigenfunction of the appropriate free-particle problem and an eigenfunction of the Schrödinger equation for a system composed of the remaining particles. Thus the type B eigenfunctions can be said to describe dissociated or ionized states of the system, although in special cases an integral combination of such functions can describe a weakly quantized state in which the system is *not* dissociated or ionized.

In order to justify the assumption of an expansion theorem bringing in the continuous spectrum we again resort to a modified problem which has no continuous spectrum but of which the actual problem can be regarded as the limit. Let F_0 denote the original problem of Eq. (32·1) and let $F_1(\sigma)$ denote a continuous one-parameter set of modified problems differing from F_0 only in that the fundamental region of coordinate space over which its solutions are spread is bounded externally by the surfaces

$$r_1 = \sigma, \qquad r_2 = \sigma, \qquad r_3 = \sigma, \qquad \cdots, \tag{32·23}$$

on which ψ is required to vanish. This problem has a system of discrete orthogonal eigenfunctions which are readily proved complete.[1]

If the radius σ is now allowed to become infinite, the problem $F_1(\sigma)$ approaches F_0 as a limit. Hence we can infer the properties of the latter spectrum from those of the former as in Sec. 30. The essential assumptions here are those which relate to orthogonality and completeness.

Both the discrete and continuous-spectrum eigenfunctions of F_0 will be degenerate as in the simple special case of the hydrogen atom. Introducing a subscript k to differentiate between different orthogonal eigenfunctions of a given eigenvalue, we designate the discrete eigenfunctions by $\psi_{kn}(x)$ and the type B functions by $\psi_k(E,x)$. We assume the quadratic integrability of the eigendifferentials

$$\Delta_j\psi_k = \int_{E_j}^{E_j+\eta_j}\psi_k(E,x)dE \tag{32.24}$$

and adopt the normalization rules

$$(\psi_{kn},\psi_{kn}) = 1; \qquad \frac{1}{\eta_j}(\Delta_j\psi_k,\Delta_j\psi_k) = 1. \tag{32.25}$$

The orthogonality of the ψ_{kn}'s was proved in Sec. 32h. Comparison with the F_1 problem suggests the additional orthogonality properties

$$\left.\begin{array}{r}(\Delta_j\psi_k,\psi_{k'n}) = 0, \\ (\Delta_j\psi_k,\Delta_l\psi_{k'}) = 0, \qquad \text{if } k \neq k', \\ (\Delta_j\psi_k,\Delta_l\psi_k) = 0, \qquad \text{if } E_j + \eta_j < E_l \text{ or } E_l + \eta_l < E_j.\end{array}\right\} \tag{32.26}$$

Finally we postulate the validity of the completeness relation

$$(f,g) = \sum_k\sum_n c_{kn}b_{kn}{}^* + \sum_k\int c_k(E)b_k(E)^*dE \tag{32.27}$$

for any pair of quadratically integrable functions f and g with Fourier coefficients c_{kn}, $c_k(E)$ and b_{kn}, $b_k(E)$, respectively, defined as in (30·43) and (30·44).

---

[1] The essential feature of the proof is to show that if we denote by $\varphi_0, \varphi_1, \cdots$ the successive normalized eigenfunctions of a sequence obtained by minimizing Q/N with orthogonality conditions similar to those used in Sec. 32h and by $E_0(\sigma), E_1(\sigma), \cdots$ the corresponding eigenvalues,

$$\lim_{n\to\infty} E_n(\sigma) = \infty.$$

This can be done with the aid of a second modification of the problem, $F_2(\sigma)$, in which the mutual repulsions of the electrons are omitted and the variables separated. Each energy level of the $F_1(\sigma)$ problem lies above the corresponding level of the $F_2(\sigma)$ problem, and the spectrum of the latter problem can be worked out explicitly and thus shown to extend to infinity. The maximum-minimum principle stated in Courant-Hilbert, *M.M.P.*, pp. 352–353, is useful in showing that the energy levels of F_1 lie above those of F_2 (*cf.* also Courant-Hilbert, *M.M.P.*, Satz 7, seite 357).

If the function $f(x)$ conforms to the class D conditions, Hf will be quadratically integrable, and in analogy with the Weyl theory we may suppose that $f(x)$ is not only root-mean-square expansible, as indicated by the completeness relation, but actually given point by point by the uniformly convergent series

$$f(x) = \sum_k \sum_n c_{kn}\psi_{kn}(x) + \sum_k \int c_k(E)\psi_k(E,x)dE. \qquad (32\cdot28)$$

We postulate that the eigendifferentials belong to the class of functions with respect to which the Hamiltonian is Hermitian. The legitimacy of a term-by-term application of the operator H to the series (32·28) is a consequence of this assumption. To prove the property, we note that as Hf belongs to class D we can expand it like $f(x)$ itself. Let the coefficients be h_{kn} and $h_k(E)$. Then

$$h_{kn} = (Hf,\psi_{kn}) = (f,H\psi_{kn}) = E_n c_{kn}; \qquad (32\cdot29)$$

$$h_k(E) = \lim_{\eta\to0} (Hf, \int_E^{E+\eta}\psi_k(E',x)dE') = \lim_{\eta\to0} (f, \int_E^{E+\eta}E'\psi_k(E',x)dE')$$

$$= Ec_k(E). \qquad (32\cdot30)$$

As these are exactly the coefficients we should obtain by applying H term by term to the series (32·28), the proposition is proved. The hypothesis on which it is based is supported by comparison of the F_0 and $F_1(\sigma)$ problems.

32k. Degeneracy.—In the relation $\lim\limits_{n\to\infty} E_n(\sigma) = \infty$ of footnote 1, p. 216, the index n is the ordinal number of the corresponding member of the complete sequence of minimizing functions $\varphi_0, \varphi_1, \cdots, \varphi_n, \cdots$. It follows that the number of linearly independent eigenfunctions for all energies below any given energy is finite. Hence each individual level of the $F_1(\sigma)$ problem must have at most a finite number of linearly independent eigenfunctions. In other words every level has a finite multiplicity or degeneracy. The question now arises, Does the multiplicity of a discrete energy level remain finite when we pass from the problem $F_1(\sigma)$ to F_0 by allowing σ to become infinite? To show that the answer is affirmative, we make the contrary assumption that one or more of the discrete eigenvalues of the F_0 problem have an infinite multiplicity. Then, since every discrete eigenfunction ψ_n of F_0 is the limit of a corresponding eigenfunction of $F_1(\sigma)$, it is evident that all eigenvalues $E_n(\sigma)$ for which n is greater than some finite value N, must approach the lowest infinitely degenerate level of F_0 as a limit when σ becomes infinite. It is further required that the complete spectrum of $F_1(\sigma)$ for large values of σ shall approach the complete spectrum of F_0 in such fashion that the spacing of the levels of $F_1(\sigma)$ shall become very small in the neighborhood

of the eigenvalues of F_0 but very large in regions where F_0 has an empty spectrum. If one considers the behavior of the spacing of a group of eigenvalues of $F_1(\sigma)$ as a function of σ in the light of these requirements, it becomes evident that the hypotheses are incompatible with the continuity of the functions $\dfrac{d}{d\sigma}E_n(\sigma)$. *Hence the discrete energy levels of the many-particle problem have at most a finite degeneracy.*

CHAPTER VII

DYNAMICAL VARIABLES AND OPERATORS

33. THE MEAN VALUES OF THE CARTESIAN COORDINATES AND CONJUGATE LINEAR MOMENTA

33a. The Statistical Mean Values of the Coordinates.—The entire theory developed in this book is based on the interpretation of $|\Psi|^2$ as probability density in the configuration space of the Cartesian positional coordinates x_1, x_2, $\cdots$, x_{3n}. In order to give this interpretation operational meaning (*cf.* footnote 1, p. 52) it is necessary to assume that the configurations of atomic systems are in principle measurable with any desired finite precision. This hypothesis is open to criticism on two counts. In the first place it overlooks the impossibility of distinguishing between different electrons, different protons, etc. In the second place it overlooks the relativistic difficulty[1] in locating the position of an electron with an uncertainty less than the Compton wave length $h/\mu_0 c$. The first of these objections is remediable by a suitable modification of the theory discussed in Sec. 42b. The second is one which has not been remedied in an entirely satisfactory manner as yet. It is apparently of slight importance so long as we restrict the application of the theory to domains in which the local wave length is everywhere large compared with the Compton wave length. In practice this means that we can apply the theory with confidence to extranuclear problems which do not involve energies for individual photons or electrons of more than, say, 100,000 electron-volts.

Setting aside both of these objections for the present, we give precise meaning to the phrase "probability density" in the following manner. Let us suppose that exact measurements of configuration are carried out on the individual members of an assemblage of N identical independent atomic systems so prepared that, at the time of measurement t_0, all are in a common subjective state described by the normalized wave function $\Psi(x,t_0) = \psi(x)$. Then, if N is large enough so that the number of systems dN found to have configurations in the element $d\tau$ of configuration space is itself large, we postulate that dN/N is equal to $|\psi|^2 d\tau$ with an error which will nearly always be small compared with unity and which can be neglected in practice.

Let q denote any function of the basic coordinates x_1, $\cdots$, x_{3n}. As the value of q is fully determined by a configuration measurement, we

---

[1] *Cf.* L. LANDAU and R. PEIERLS, *Zeits. f. Physik* **69**, 56 (1931).

can work out the probability of any range dq of q values, say $q' < q < q''$, by integrating $|\Psi|^2 d\tau$ over that portion of configuration space for which q is in the range dq. Thus our hypothesis regarding $|\Psi|^2$ implicitly fixes the distribution function for measured values of q made on an assemblage of systems in the state Ψ. Moreover, it fixes the mean value[1] of any such function $q(x_1, \cdots, x_{3n})$ for a sufficiently large number of measurements as

$$\bar{q} = \int_\infty q|\Psi|^2 d\tau = (q\Psi, \Psi). \tag{33·1}$$

This type of statistical mean value is often referred to as the *expectation value* of the quantity in question. Such mean values were used in Sec. 13 in deriving Newton's second law of motion for wave packets.

In order to compute the mean values of other physical quantities, involving the measurement of velocities as well as position, a more difficult computation is generally necessary, but with the aid of suitably defined operators it is possible to set up formulas for mean values which are very similar to (33·1). The way in which this is done will be illustrated in Sec. 33b.

33b. The Linear Momentum Operator.—In Chap. II, Sec. 15, we saw that, by Fourier analysis of $\Psi(x,y,z,t)$, we can derive a probability amplitude $\Phi(p_x,p_y,p_z,t)$ for the linear momentum of a particle, or system of particles, such that $\Phi\Phi^* dp_x dp_y dp_z$ gives the probability that the momentum vector terminates in the volume element $dp_x dp_y dp_z$ of momentum space. With the aid of Φ we can compute the mean value of any component of the linear momentum, say p_x, by

$$\bar{p}_x = \iiint_\infty \Phi^* p_x \Phi \, dp_x dp_y dp_z = (p_x\Phi, \Phi). \tag{33·2}$$

An alternative mode of averaging, which does not involve the evaluation of the Φ function, can be derived by means of the operator $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial x}$. When applied to a wave function corresponding to a unique value of p_x, and hence of the form $\Psi = f(y,z,t)e^{2\pi i x p_x/h}$, this operator yields the relation

$$\left(\frac{h}{2\pi i}\frac{\partial}{\partial x}\right)\Psi = p_x\Psi. \tag{33·3}$$

Hence, differentiating the general formula of Eq. (15·9) with respect to x, we obtain

---

[1] As emphasized by VON NEUMANN [*Göttinger Nachrichten, Math.-phys. Klasse,* 248 (1928)], a knowledge of the mean values of all functions of the coordinates is equivalent to a knowledge of their distribution function. In other words, the validity of (33·1) for every q is a necessary and sufficient condition for the validity of the statement that $|\Psi|^2$ is a probability density in configuration space. Hence (33·1) is sometimes taken as the basic postulate rather than our assumption regarding $|\Psi|^2$.

$$\frac{h}{2\pi i}\frac{\partial\Psi}{\partial x} = h^{-3/2}\int\int\int_{\infty} p_x\Phi e^{\frac{2\pi i}{h}(xp_x+yp_y+zp_z)}dp_x dp_y dp_z. \qquad (33\cdot4)$$

It will be observed that the right-hand member of the above equation gives the Fourier analysis of the function $\dfrac{h}{2\pi i}\dfrac{\partial\Psi}{\partial x}$ into plane harmonic waves. In other words, $p_x\Phi$ is the *Fourier transform* of $\dfrac{h}{2\pi i}\dfrac{\partial\Psi}{\partial x}$, just as Φ is the Fourier transform of Ψ. But by a known theorem of Fourier analysis[1] the scalar product of two quadratically integrable functions is equal to the scalar product of their Fourier transforms.

Hence

$$\left(\frac{h}{2\pi i}\frac{\partial\Psi}{\partial x},\Psi\right) = (p_x\Phi,\Phi),$$

or

$$\overline{p_x} = \left(\frac{h}{2\pi i}\frac{\partial\Psi}{\partial x},\Psi\right) = \int\int\int_{\infty}\Psi^*\left(\frac{h}{2\pi i}\frac{\partial}{\partial x}\right)\Psi dx dy dz. \qquad (33\cdot5)$$

Thus the mean value of each of the components of linear momentum can be evaluated by a rule formally the same as that used for $f(q)$, but with the substitution of the operator $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial q}$ for the momentum component to be averaged. We might say that the average value of the momentum p_x in momentum space is equal to the "average value of the corresponding operator $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial x}$ in x,y,z space."

In the same way we can formally determine the mean value of any positive power of p by the formula

$$\overline{p^n} = \int\int\int_{\infty}\Psi^*\left(\frac{h}{2\pi i}\frac{\partial}{\partial q}\right)^n\Psi dx dy dz. \qquad (33\cdot6)$$

From Eqs. (33·6) and (5·5) it follows that the mean value of the square of the total linear momentum for a single-energy wave function in three dimensions is equal to the mean value of the square of the classical local momentum. Thus

$$\overline{p^2} = \overline{p_x{}^2} + \overline{p_y{}^2} + \overline{p_z{}^2} = \overline{2\mu(E-V)}. \qquad (33\cdot7)$$

Using this definition of the statistical mean value of a component of linear momentum we observe that Eqs. (13·8) and (13·9) are equivalent to

$$\mu\frac{d\bar{x}}{dt} = \overline{p_x}; \qquad \frac{d\overline{p_x}}{dt} = -\overline{\frac{\partial V}{\partial x}}. \qquad (33\cdot8)$$

[1] *Cf.* footnote 1, p. 36; also Eqs. (30·29) and (32·27).

As Sommerfeld,[1] for instance, has pointed out, the use of the operator $\frac{h}{2\pi i}\frac{\partial}{\partial q_k}$ for the momentum p_k conjugate to the coordinate q_k is intimately related to a fundamental theorem derived by Schrödinger[2] and interpreted initially as a statement of the law of the conservation of electricity. In its simple form for a single charged particle in three dimensions, using Cartesian coordinates, this theorem [*cf.* Eqs. (8·5) and (8·6)] is

$$\frac{\partial}{\partial t}\Psi\Psi^* = \frac{h}{4\pi\mu i}\,\mathrm{div}\left[\Psi\,\mathrm{grad}\,\Psi^* - \Psi^*\,\mathrm{grad}\,\Psi + \frac{4\pi i e}{hc}\vec{a}\Psi\Psi^*\right]. \quad (33\cdot9)$$

Here $\vec{a}$ is the vector potential of an assumed external electromagnetic field. If we apply this theorem to the wave function for an electron of charge e, we may interpret $e\Psi\Psi^*$ as the statistical mean charge density. Equation (33·9) then takes the form of the equation of continuity in hydrodynamics with the vector density of electric current defined by[3]

$$\frac{eh}{4\pi\mu i}\left[\Psi^*\,\mathrm{grad}\,\Psi - \Psi\,\mathrm{grad}\,\Psi^* - \frac{4\pi i e}{hc}\vec{a}\Psi\Psi^*\right] = \frac{e}{\mu}\vec{I}. \quad (33\cdot10)$$

Multiplication by μ/e gives the mass current density $\vec{I}$, and integration over all space should give the product of the mass μ into the average vector velocity. Therefore, if we define the momentum in an electromagnetic field by Eq. (15·24), we obtain for its average x component, for example, the value

$$\overline{p_x} = \mu\dot{x} + \frac{e}{c}\overline{a_x} = \frac{h}{4\pi i}\int_\infty\left[\Psi^*\frac{\partial\Psi}{\partial x} - \Psi\frac{\partial\Psi^*}{\partial x}\right]d\tau. \quad (33\cdot11)$$

But the operator $\frac{h}{2\pi i}\frac{\partial}{\partial x}$ is clearly Hermitian with respect to configuration space and the class of all continuous and piece-by-piece differentiable functions which approach zero at infinity. This includes class D. Hence Eq. (33·11) is equivalent to (33·5), and the above discussion yields a new derivation of the latter formula independent of the Fourier integral theorem.

Another important application of the operator $\frac{h}{2\pi i}\frac{\partial}{\partial q}$ is the general proof of the Heisenberg inequality $\Delta p\Delta q \geqslant h/4\pi$ for linear momentum (*cf.* Chap. II, Sec. 16). To establish this inequality we first transform the formula for $(\Delta p_k)^2$ given in (16·1) to

$$(\Delta p_k)^2 = \int_\infty\Psi^*\left(\frac{h}{2\pi i}\frac{\partial}{\partial q_k} - \overline{p_k}\right)^2\Psi d\tau \quad (33\cdot12)$$

[1] SOMMERFELD, *Atombau und Spektrallinien, Wellenmechanischer Ergänzungsband,* pp. 284–285, Braunschweig, 1929.
[2] E. SCHRÖDINGER, *Ann. d. Physik* (4) **81**, 136 (1926).
[3] *Cf.* CONDON and MORSE, *Q.M.,* p. 28.

by means of (33·5).   Following Weyl[1] we use a slight generalization of the Schwarz inequality [cf. Eqs. (22·10) and (22·20)] which states that if f_1, g_1, f_2, g_2 are any four quadratically integrable functions, their scalar products are subject to the inequality

$$[(f_1, f_1) + (f_2, f_2)][(g_1, g_1) + (g_2, g_2)] \geqq |(f_1, g_1) + (f_2, g_2)|^2. \quad (33·13)$$

Making the identifications

$$f_1 = \left(\frac{h}{2\pi i}\frac{\partial}{\partial q_k} - \overline{p_k}\right)\Psi = f_2{}^*,$$
$$g_1 = (q_k - \overline{q_k})\Psi \qquad = -g_2{}^*,$$

and using the Hermitian character of the differential operator $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial q_k}$ and the multiplication operator $[q_k \times]$ with respect to coordinate space and class D functions, we readily deduce the relations

$$(f_1, f_1) + (f_2, f_2) = 2\int\left|\left(\frac{h}{2\pi i}\frac{\partial}{\partial q_k} - \overline{p_k}\right)\Psi\right|^2 d\tau = 2\Delta p_k{}^2, \quad (33·14)$$

$$(g_1, g_1) + (g_2, g_2) = 2\Delta q_k{}^2, \quad (33·15)$$

when Ψ is of class D.   Similarly,

$$(f_1, g_1) + (f_2, g_2) = \int \Psi^*\left[q_k\frac{h}{2\pi i}\frac{\partial}{\partial q_k} - \frac{h}{2\pi i}\frac{\partial}{\partial q_k}q_k\right]\Psi d\tau. \quad (33·16)$$

The integrand of the right-hand· member of (33·16) reduces at once to $-\dfrac{h\Psi\Psi^*}{2\pi i}$ [cf. Eq. (37·5)] and the integral itself reduces to $-\dfrac{h}{2\pi i}$.   Combining this result with (33·14), (33·15), and (33·13), we obtain

$$\Delta p_k{}^2\Delta q_k{}^2 \geqq \left(\frac{h}{4\pi}\right)^2, \quad (33·17)$$

which is the square of the desired Heisenberg inequality.

The machinery at hand permits a useful generalization of our previous treatment of the variation of $\Delta q_k{}^2$ in time.   For our purposes it will suffice to consider the free particle for which $\Delta p_k{}^2$ is constant.   We shall carry out the calculation for the one-dimensional case, the extension to three dimensions being obvious.

The momentum probability amplitude is given by

$$\Phi = h^{\frac{1}{2}}\int_{-\infty}^{\infty} \Psi e^{-\frac{2\pi i q p}{h}} dq.$$

Hence

$$-\frac{h}{2\pi i}\frac{\partial\Phi}{\partial p} = h^{-\frac{1}{2}}\int_{-\infty}^{\infty} q\Psi e^{-\frac{2\pi i q p}{h}} dq.$$

[1] HERMANN WEYL, *The Theory of Groups and Quantum Mechanics*, either 1st or 2d ed., Appendix I.   For other proofs see W. Pauli, Jr., in Geiger and Scheel's *Handbuch der Physik*, XXIV/1, 2d ed., p. 102, Berlin, 1933; also W. Heisenberg, *The Physical Principles of the Quantum Theory*, pp. 15–19, Chicago, 1930.

By Plancherel's theorem (*cf.* footnote 3, p. 36),

$$\bar{q} = \int_{-\infty}^{\infty} \Psi^* q \Psi dq = \int_{-\infty}^{\infty} \Phi^* \left[-\frac{h}{2\pi i} \frac{\partial \Phi}{\partial p} \right] dp;$$

$$\overline{q^2} = \int_{-\infty}^{\infty} \Psi^* q^2 \Psi dq = \int_{-\infty}^{\infty} \left[\frac{h}{2\pi i} \frac{\partial \Phi^*}{\partial p} \right] \left[-\frac{h}{2\pi i} \frac{\partial \Phi}{\partial p} \right] dp.$$

Now for the free particle

$$\Phi(p,t) = e^{-\frac{i\pi p^2 (t - t_0)}{\mu h}} \Phi_0, \qquad \Phi_0 \equiv \Phi(p,t_0);$$

$$-\frac{h}{2\pi i} \frac{\partial \Phi}{\partial p} = \frac{p(t - t_0)}{\mu} \Phi + e^{-\frac{i\pi p^2 (t - t_0)}{\mu h}} \left[-\frac{h}{2\pi i} \frac{\partial \Phi_0}{\partial p} \right].$$

Hence

$$\bar{q} = \frac{(t - t_0)}{\mu} \int_{-\infty}^{\infty} \Phi^* p \Phi dp + \int_{-\infty}^{\infty} \Phi_0^* \left(-\frac{h}{2\pi i} \frac{\partial \Phi_0}{\partial p} \right) dp \equiv \frac{\bar{p}(t - t_0)}{\mu} + \bar{q}_0;$$

$$\overline{q^2} = \frac{(t - t_0)^2}{\mu^2} \int_{-\infty}^{\infty} \Phi^* p^2 \Phi dp + \int_{-\infty}^{\infty} \left(\frac{h}{2\pi i} \frac{\partial \Phi_0^*}{\partial p} \right) \left(-\frac{h}{2\pi i} \frac{\partial \Phi_0}{\partial p} \right) dp$$

$$+ \frac{t - t_0}{\mu} \int_{-\infty}^{\infty} p \left[\Phi_0^* \left(-\frac{h}{2\pi i} \frac{\partial \Phi_0}{\partial p} \right) + \Phi_0 \left(\frac{h}{2\pi i} \frac{\partial \Phi_0^*}{\partial p} \right) \right] dp$$

$$= \overline{p^2} \frac{(t - t_0)^2}{\mu^2} + (\overline{q^2})_{t-t_0} + \frac{t - t_0}{\mu} (\overline{pq} + \overline{qp})_{t-t_0}.$$

Since $\Delta q^2 = \overline{q^2} - \bar{q}^2$ and $\Delta p^2 = \overline{p^2} - \bar{p}^2$, we have

$$(\Delta q^2)_t = (\Delta q^2)_{t_0} + \frac{\Delta p^2 (t - t_0)^2}{\mu^2} + \frac{(t - t_0)}{\mu} (\overline{pq} + \overline{qp} - 2\bar{p}\bar{q})_{t_0} \qquad (33 \cdot 18)$$

It is readily verified that $(\overline{pq} + \overline{qp} - 2\bar{p}\bar{q})_{t-t_0}$ vanishes for the special cases of Secs. 9 and 16. It must do this in the latter case since the presence of a term linear in the time in that case would lead to a violation of the uncertainty principle. In any case Δq^2 has a minimum at some time t', so that by choosing $t_0 = t'$ we have simply

$$(\Delta q^2)_t = (\Delta q^2)_{t_0} + \frac{\Delta p^2 (t - t_0)^2}{\mu^2}.$$

For further discussion, see W. Pauli, in Geiger and Scheel's *Handbuch der Physik*, XXIV/1, 2d ed., p. 100, Berlin, 1933. Here the term linear in t is interpreted in terms of the probability current density.

34. THE ANGULAR-MOMENTUM OPERATORS

34a. Definition of Operators.—In the classical mechanics the angular momentum of a system of particles is defined by the formula

$$\vec{\mathcal{L}} = \sum_{k=1}^{n} \vec{r}_k \times \vec{p}_k, \qquad (34 \cdot 1)$$

where $\vec{p}_k$ is the linear momentum of the kth particle and $\vec{r}_k$ is its distance from the origin. With the aid of the corresponding classical expressions for the components of $\vec{\mathcal{L}}$, it is easy to invent operators which have the

same relation to the components of angular momentum as $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial x}$ has to the x component of linear momentum.

Thus, in the case of a single particle where $\mathcal{L}_z = xp_y - yp_x$, we replace p_x, p_y by the corresponding operators to obtain an operator for $\mathcal{L}_z$ which we designate for the present as $(\mathcal{L}_z)_{op}$

$$(\mathcal{L}_z)_{op} = \frac{h}{2\pi i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right). \tag{34·2}$$

In harmony with Eq. (33·3) let us adopt the convention that a particle whose state is described by Ψ shall be said to have a unique value of $\mathcal{L}_z$, say $\mathcal{L}_z'$, if the application of the operator $(\mathcal{L}_z)_{op}$ to Ψ is equivalent to multiplication by $\mathcal{L}_z'$. In other words, $\mathcal{L}_z$ has the unique value $\mathcal{L}_z'$ if Ψ is a solution of the differential equation

$$(\mathcal{L}_z)_{op}\Psi \equiv \frac{h}{2\pi i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\Psi = \mathcal{L}_z'\Psi. \tag{34·3}$$

Otherwise we assign no definite value of $\mathcal{L}_z$ to the particle under consideration but assume that it has a certain probability of taking on any one of a number of values, like the variety of values of p_x, p_y, p_z in the case of a wave packet. As a means for determining the probability amplitude for $\mathcal{L}_z$ in such cases we adopt the scheme of analyzing the wave function into a linear combination of functions which satisfy Eq. (34·3) together with suitable boundary and continuity conditions for some value of the parameter $\mathcal{L}_z'$. Thus Eq. (34·3) is made the basis of an eigenvalue-eigenfunction problem like the Schrödinger equation (5·9). The eigenvalues of (34·3) are the possible values of $\mathcal{L}_z$ and the eigenfunctions describe states having definite values of $\mathcal{L}_z$.

34b. Hermitian Character of Angular-momentum Operators.—By direct application of Gauss's transformation the operator $(\mathcal{L}_z)_{op}$ is readily proved to be Hermitian with respect to the class of all quadratically integrable functions which have quadratically integrable transforms by $(\mathcal{L}_z)_{op}$. Its Hermitian character can also be deduced from the Hermitian character of $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial x}$ and $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial y}$ by means of the relations

$$\left(x\frac{h}{2\pi i}\frac{\partial}{\partial y}\psi_1,\ \psi_2\right) = \left(\frac{h}{2\pi i}\frac{\partial}{\partial y}\psi_1,\ x\psi_2\right) = \left(\psi_1,\ x\frac{h}{2\pi i}\frac{\partial}{\partial y}\psi_2\right),$$

$$\left(y\frac{h}{2\pi i}\frac{\partial}{\partial x}\psi_1,\ \psi_2\right) = \left(\frac{h}{2\pi i}\frac{\partial}{\partial x}\psi_1,\ y\psi_2\right) = \left(\psi_1,\ y\frac{h}{2\pi i}\frac{\partial}{\partial x}\psi_2\right).$$

$(\mathcal{L}_x)_{op}$ and $(\mathcal{L}_y)_{op}$ can be treated in the same way.

Since these operators are Hermitian, their quadratically integrable eigenfunctions have real eigenvalues. The eigenfunctions of this type

for different eigenvalues are mutually orthogonal. Since the existence of a continuous spectrum of eigenvalues necessarily involves the existence of a continuous array of eigenfunctions depending on the eigenvalue parameter, we can readily prove that quadratically integrable eigenfunctions of an Hermitian operator O cannot be associated with a continuous spectrum. For if such a continuous array of eigenfunctions, say $\psi(\lambda,x)$, did exist, the scalar product $(\psi(\lambda,x),\ \psi(\lambda',x))$ would be a continuous function of λ and λ' which vanishes when $\lambda \neq \lambda'$ but does not vanish when $\lambda = \lambda'$—an obvious impossibility.[1]

34c. The Expansion Theorem.—In order to justify our definition of the angular-momentum operators we must develop an expansion theorem for their eigenfunctions. In this particular case the simplest procedure is to begin by finding the actual form of the eigenfunctions in spherical coordinates. The expansion theorem then turns out to be a simple Fourier expansion.

Direct transformation to the coordinates r, θ, φ of Sec. 28 carries the operator $x\dfrac{\partial}{\partial y} - y\dfrac{\partial}{\partial x}$ over into $\dfrac{\partial}{\partial \varphi}$, and Eq. (34·3) becomes

$$\frac{h}{2\pi i}\frac{\partial \Psi}{\partial \varphi} = \mathcal{L}_z'\Psi. \tag{34·4}$$

It has the solutions

$$\Psi = \chi(r,\theta,t)e^{\frac{2\pi i \mathcal{L}_z'\varphi}{h}}. \tag{34·5}$$

By the reasoning of Sec. 28 Ψ must have the period 2π in the argument φ if Ψ is to be a continuous single-valued function of x, y, z. Hence the quantity $2\pi\mathcal{L}_z'/h$ is restricted to integral eigenvalues.

$$\mathcal{L}_z' = \frac{mh}{2\pi}. \qquad m = 0,\ \pm 1,\ \pm 2,\ \cdots \tag{34·6}$$

The eigenfunctions can all be made quadratically integrable by suitable choice of the arbitrary factor χ. In accordance with the general theorem stated above (Sec. 34b), we have a discrete spectrum only for the "physical quantity" $\mathcal{L}_z$. From our present point of view nonintegral values of $2\pi\mathcal{L}_z'/h$ have no meaning.

Since an arbitrary single-valued continuous function of the Cartesian coordinates transforms in spherical coordinates into a function with the period 2π in the azimuthal angle φ, any class A or class D wave function for a single particle can be analyzed into the complex Fourier's series of $\mathcal{L}_z$ eigenfunctions

$$\Psi = \sum_{m=-\infty}^{+\infty} \chi_m(r,\theta,t)e^{im\varphi}. \tag{34·7}$$

[1] *Cf.* J. Frenkel, *Wellenmechanik.* Chap. III, Sec. 3, p. 152.

The coefficients $\chi_m(r,\theta,t)$ are determined by the usual formula

$$\chi_m = \frac{1}{2\pi}\int_0^{2\pi}\Psi e^{-im\varphi}d\varphi,$$

and satisfy the completeness relation

$$\int_0^{2\pi}\Psi\Psi^*d\varphi = 2\pi\sum_{m=-\infty}^{+\infty}\chi_m\chi_m^* \tag{34·8}$$

[*cf.* Eq. (25·11)]. It follows that if Ψ is normalized,

$$2\pi\sum_{m=-\infty}^{+\infty}\int_0^\infty r^2dr\int_0^\pi \sin\theta\,|\chi_m|^2d\theta = \iiint_\infty|\Psi|^2dxdydz = 1. \tag{34·9}$$

The contribution of the eigenfunction $\chi_m e^{im\varphi}$ to the integrated intensity of the wave system is $2\pi\iint r^2 \sin\theta\,|\chi_m|^2drd\theta$. We infer that the series of functions $(2\pi)^{1/2}\chi_m(r,\theta,t)$ play the role of probability amplitude for the set of independent variables r, θ, $\mathfrak{L}_z{}'$ (*cf.* Sec. 15, p. 63).

By shifting the axes of our spherical coordinates we obtain the eigenfunctions for $\mathfrak{L}_x$ and $\mathfrak{L}_y$. Their eigenvalues are of course the same as those of $\mathfrak{L}_z$.

34d. Angular Momentum of a System of Particles.—In the general case of a system of $n = f + 1$ particles we define the operators for the three components of angular momentum with respect to the origin of an absolute coordinate system as

$$\left.\begin{aligned}
(\mathfrak{L}_x)_{op} &= \frac{h}{2\pi i}\sum_{k=1}^{f+1}\left(y_k\frac{\partial}{\partial z_k} - z_k\frac{\partial}{\partial y_k}\right), \\
(\mathfrak{L}_y)_{op} &= \frac{h}{2\pi i}\sum_{k=1}^{f+1}\left(z_k\frac{\partial}{\partial x_k} - x_k\frac{\partial}{\partial z_k}\right), \\
(\mathfrak{L}_z)_{op} &= \frac{h}{2\pi i}\sum_{k=1}^{f+1}\left(x_k\frac{\partial}{\partial y_k} - y_k\frac{\partial}{\partial x_k}\right).
\end{aligned}\right\} \tag{34·10}$$

These operators are also Hermitian with respect to coordinate space and the class of single-valued differentiable functions which are quadratically integrable and have quadratically integrable transforms.

In order to resolve them into constituents corresponding to the internal angular momentum, and to the angular momentum of the center of gravity with respect to the origin, we make the change of variables given in Eqs. (15·13) and (15·14). The expression for $(\mathfrak{L}_z)_{op}$ becomes

$$(\mathfrak{L}_z)_{op} = \frac{h}{2\pi i}\left[\left(X\frac{\partial}{\partial Y} - Y\frac{\partial}{\partial X}\right) + \sum_{k=1}^{f}\left(\xi_k\frac{\partial}{\partial \eta_k} - \eta_k\frac{\partial}{\partial \xi_k}\right)\right]. \tag{34·11}$$

We define the operator for the z component of the *internal* angular momentum to be $\dfrac{h}{2\pi i}\displaystyle\sum_{k=1}^{f}\left(\xi_k\dfrac{\partial}{\partial\eta_k} - \eta_k\dfrac{\partial}{\partial\xi_k}\right)$. Since the same form of operator is used for a given component of the internal angular momentum of a system of $f + 1$ particles and for the absolute angular momentum of a system of f particles, the eigenvalues for the two cases must be the same and the eigenfunctions must have the same form.

In order to determine the spectrum of $(\mathcal{L}_z)_{op}$ we revert to the last of the formulas (34·10) and replace the Cartesian coordinates of the various particles by spherical coordinates. $(\mathcal{L}_z)_{op}$ is reduced to the form

$$(\mathcal{L}_z)_{op} = \frac{h}{2\pi i}\sum_{k=1}^{f}\frac{\partial}{\partial\varphi_k}. \qquad (34\cdot12)$$

Introducing the relative azimuthal coordinates[1]

$$\varphi = \varphi_1, \qquad \alpha_2 = \varphi_2 - \varphi_1, \qquad \alpha_3 = \varphi_3 - \varphi_1, \qquad \cdots \qquad \alpha_f = \varphi_f - \varphi_1,$$

we finally obtain

$$(\mathcal{L}_z)_{op} = \frac{h}{2\pi i}\frac{\partial}{\partial\varphi}. \qquad (34\cdot13)$$

Thus the equation defining the eigenfunctions and eigenvalues of $(\mathcal{L}_z)_{op}$ is identical in form with that used for a single particle. The eigenvalues are again the integral multiples of $h/2\pi$ and the eigenfunctions are

$$\Psi_M = \chi_M(r_1,\theta_1,r_2,\theta_2,\alpha_2,\ \cdots\ )e^{iM\varphi}. \qquad M = 0, \pm1, \pm2, \cdots \quad (34\cdot14)$$

Here we adopt the usual practice of spectroscopists in using a capital M as the quantum number associated with $\mathcal{L}_z$ when referring to the resultant angular momentum of several particles.

Comparing (34·14) with (28·10) we see that the factorable solutions of the two-particle problem with the center of gravity eliminated are eigenfunctions of $(\mathcal{L}_z)_{op}$ provided that the factor $\Phi(\varphi)$ has the exponential form $e^{iM\varphi}$. The trigonometric forms of Eq. (28·10) are linear combinations of $e^{iM\varphi}$ and $e^{-iM\varphi}$ and as such have definite values of $\mathcal{L}_z{}^2$. In fact Eq. (28·9) can be set up by forming the operator for $\mathcal{L}_z{}^2$, *viz.*, $-\left(\dfrac{h}{2\pi}\right)^2\dfrac{\partial^2}{\partial\varphi^2}$, and requiring that the application of this operator to Ψ shall yield a multiple of Ψ.

---

[1] The coordinate system

$$\varphi = \frac{1}{f}\sum_{k=1}^{f}\varphi_k, \qquad \alpha_2 = \varphi_2 - \varphi_1, \qquad \alpha_3 = \varphi_3 - \varphi_2, \qquad \cdots \qquad \alpha_f = \varphi_f - \varphi_{f-1}$$

will do as well and is more symmetrical.

The expansion theorem of (34·7) holds equally well for the many-particle problem and for the single-particle problem except that the coefficients in the former case depend on a larger group of independent variables. The contribution of the eigenfunction correlated with M to the integrated intensity of the wave system is

$$2\pi \int \cdots \int \chi_M \chi_M{}^* \prod_{k=1}^{f} r_k^2 \sin \theta_k \, dr_k d\theta_k \prod_{l=2}^{f} d\alpha_l.$$

We interpret this contribution as the probability that an appropriate measurement of $\mathcal{L}_z$ will yield the value $Mh/2\pi$. This statement pre-. scribes in a general way the experimental procedure which must be used in determining the distribution of $\mathcal{L}_z$ values for any given assemblage of systems, *viz.*, the experimental arrangement must be such as to resolve the primary assemblage into subassemblages each of which has a wave function of the form $\chi_M e^{iM\varphi}$. In the Stern-Gerlach experiment, which measures both atomic magnetic moments and $\mathcal{L}_z$ values, this resolution is accomplished by sending a beam of atoms through an inhomogeneous magnetic field directed along the z axis in which the path of the atoms in each subassemblage is different from that of atoms in any other subassemblage.

34e. Mean Values.—To get the mean, or expectation, value of $\mathcal{L}_z$ for any given wave function we use the same procedure as for one of the components of linear momentum.

$$\begin{aligned}
\mathcal{L}_z &= 2\pi \sum_{M=-\infty}^{+\infty} \int \cdots \int \frac{Mh}{2\pi} \chi_M \chi_M{}^* \prod_{k=1}^{f} r_k^2 \sin \theta_k \, dr_k d\theta_k \prod_{l=2}^{f} d\alpha_l \\
&= \int \cdots \int \Psi^* \frac{h}{2\pi i} \frac{\partial}{\partial \varphi} \Psi d\tau \\
&= \int \cdots \int \Psi^* \sum_k \frac{h}{2\pi i} \left(\xi_k \frac{\partial}{\partial \eta_k} - \eta_k \frac{\partial}{\partial \xi_k} \right) \Psi d\tau.
\end{aligned} \right\} \quad (34\cdot15)$$

With the aid of the above expression we can show that the definition of $\mathcal{L}_z$ given by Eq. (34·3) is equivalent to the classical definition in the limiting case of a sharply defined wave packet. For such a packet, where the range of values of the Cartesian coordinates and momenta is small compared with the corresponding mean values,[1]

[1] To prove Eq. (34·16) we note that, if Ψ is appreciably different from zero only in the immediate neighborhood of the point $x, y, z = \bar{x}, \bar{y}, \bar{z}$, we can substitute $\bar{\xi}, \bar{\eta}$ for ξ, η in Eq. (34·15) without appreciable error. Then

$$\int \cdots \int \Psi^* \sum_k \frac{h}{2\pi i} \left(\xi_k \frac{\partial}{\partial \eta_k} - \eta_k \frac{\partial}{\partial \xi_k} \right) \Psi d\tau$$

$$= \sum_k \left(\bar{\xi}_k \int \cdots \int \Psi^* \frac{h}{2\pi i} \frac{\partial}{\partial \eta_k} \Psi d\tau - \bar{\eta}_k \int \cdots \int \Psi^* \frac{h}{2\pi i} \frac{\partial}{\partial \xi_k} \Psi d\tau \right).$$

$$\mathfrak{L}_{\bar{z}} = \sum_k \int \cdots \int \Psi^* \frac{h}{2\pi i}\left(\xi_k \frac{\partial}{\partial \eta_k} - \eta_k \frac{\partial}{\partial \xi_k}\right)\Psi d\tau = \sum_k (\bar{\xi}_k \bar{p}_{\eta_k} - \bar{\eta}_k \bar{p}_{\xi_k}).$$

(34·16)

Since the classical values of coordinates and momenta are the same as the mean values for the packet, this shows that our definition is equivalent to the classical one in the realm of validity of the older theory.

34f. The Vector Angular Momentum and Its Square; the Symmetric Top.—The possibility of assigning an exact value to $\mathfrak{L}_z$ and hence also to $\mathfrak{L}_x$ or $\mathfrak{L}_y$ does not carry with it the possibility of assigning an exact value to the vector angular momentum $\overrightarrow{\mathfrak{L}}$, for there are no simultaneous nonvanishing solutions of the characteristic equations for the three components. On the other hand, unique values for the square of the total angular momentum are possible if we define this quantity by means of the natural operator

$$(\mathfrak{L}^2)_{\mathrm{op}} = [(\mathfrak{L}_x)_{\mathrm{op}}]^2 + [(\mathfrak{L}_y)_{\mathrm{op}}]^2 + [(\mathfrak{L}_z)_{\mathrm{op}}]^2$$

$$= -\frac{h^2}{4\pi^2}\left\{\left[\sum_{k=1}^{f}\left(\eta_k\frac{\partial}{\partial \zeta_k} - \zeta_k\frac{\partial}{\partial \eta_k}\right)\right]^2 + \left[\sum_{k=1}^{f}\left(\zeta_k\frac{\partial}{\partial \xi_k} - \xi_k\frac{\partial}{\partial \zeta_k}\right)\right]^2 + \left[\sum_{k=1}^{f}\left(\xi_k\frac{\partial}{\partial \eta_k} - \eta_k\frac{\partial}{\partial \xi_k}\right)\right]^2\right\}. \quad (34\cdot17)$$

$(\mathfrak{L}^2)_{\mathrm{op}}$ is readily proved Hermitian with respect to the class of single-valued differentiable functions which, together with their transforms by $(\mathfrak{L}_x)_{\mathrm{op}}, (\mathfrak{L}_y)_{\mathrm{op}}, (\mathfrak{L}_z)_{\mathrm{op}}, [(\mathfrak{L}_x)_{\mathrm{op}}]^2, [(\mathfrak{L}_y)_{\mathrm{op}}]^2, [(\mathfrak{L}_z)_{\mathrm{op}}]^2$, are quadratically integrable.

In our study of the eigenvalue-eigenfunction problem for $(\mathfrak{L}^2)_{\mathrm{op}}$ we begin with the three-dimensional case where f is unity. Introducing spherical coordinates, applying the operator to Ψ, and requiring that the result shall be equal to $(\mathfrak{L}^2)'\Psi$, if $\mathfrak{L}^2$ is unique, we obtain the differential equation

$$(\mathfrak{L}^2)_{\mathrm{op}}\Psi = -\left(\frac{h}{2\pi}\right)^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\Psi}{\partial\varphi^2}\right] = (\mathfrak{L}^2)'\Psi. \quad (34\cdot18)$$

This is identical with the differential equation (28·6) whose solutions are tesseral harmonics and whose eigenvalues are given by (28·15). Thus

$$(\mathfrak{L}^2)' = l(l+1)\left(\frac{h}{2\pi}\right)^2, \qquad l = 0, 1, 2, 3 \cdots \quad (34\cdot19)$$

confirming the interpretation of the quantity $l(l+1)(h/2\pi)^2$ which was made in Sec. 28 (p. 151).

In order to deduce the eigenfunctions and eigenvalues of $\mathcal{L}^2$ for a many-particle problem we introduce a set of coordinate axes x', y', z' so chosen that the z' axis passes through the first particle and the $x'z'$ plane through the second particle. The relative positions of the various particles are then fixed by the coordinates z_1', x_2', z_2', x_3', y_3', $\cdots$, z_f'. The remaining coordinates are taken to be three Eulerian angles φ, θ, ψ which define the orientation of the x', y', z' axes with respect to a set of fixed axes x, y, z. These angles are the angles of three successive rotations required to carry a set of movable coordinates ξ, η, ζ from an initial orientation in coincidence with the x, y, z system to a final orientation in coincidence with the x', y', z' system. The first rotation is made through an angle φ in the positive sense (by the right-handed screw rule) about the z axis. φ is so chosen that the η axis is carried into coincidence with the intersection of the x, y and x', y' planes (the line of nodes NN in Fig. 14). There are two possible values of φ in the range $0 \leqslant \varphi \leqslant 2\pi$ which satisfy this condition, but we resolve the ambiguity by giving φ a value less than π if, and only if, the projection of the positive z' axis on the x,y plane makes an acute angle with

FIG. 14. —Euler's angles.

the positive y axis. Then by applying a second rotation to the ξ, η, ζ axes through a suitable angle θ in the positive sense about the second position of the η axis (line of nodes NN), we can bring the ξ,η plane into coincidence with the x',y' plane and the ζ axis into coincidence with the z' axis. The above choice of the angle φ is such that θ need never exceed the value π. A third rotation through a suitable angle ψ $(0 \leqslant \psi \leqslant 2\pi)$ about the z' axis brings the ξ,η,ζ system into its final configuration. In terms of these angles the direction cosines relating the primed and unprimed sets of coordinates are as follows:

	x	y	z
x'	$\cos\varphi\cos\theta\cos\psi - \sin\varphi\sin\psi$	$\sin\varphi\cos\theta\cos\psi + \cos\varphi\sin\psi$	$-\sin\theta\cos\psi$
y'	$-\cos\varphi\cos\theta\sin\psi - \sin\varphi\cos\psi$	$-\sin\varphi\cos\theta\sin\psi + \cos\varphi\cos\psi$	$\sin\theta\sin\psi$
z'	$\cos\varphi\sin\theta$	$\sin\varphi\sin\theta$	$\cos\theta$

(34·20)

The expressions for the operators $(\mathcal{L}_x)_{op}$, $(\mathcal{L}_y)_{op}$, $(\mathcal{L}_z)_{op}$ now take the forms

$$(\mathcal{L}_x)_{\text{op}} = \frac{h}{2\pi i}\left(-\cos\,\varphi\cot\,\theta\,\frac{\partial}{\partial\varphi} - \sin\,\varphi\frac{\partial}{\partial\theta} + \frac{\cos\,\varphi}{\sin\,\theta}\,\frac{\partial}{\partial\psi}\right), \quad (34\cdot21)$$

$$(\mathcal{L}_y)_{\text{op}} = \frac{h}{2\pi i}\left(-\sin\,\varphi\cot\,\theta\,\frac{\partial}{\partial\varphi} + \cos\,\varphi\frac{\partial}{\partial\theta} + \frac{\sin\,\varphi}{\sin\,\theta}\,\frac{\partial}{\partial\psi}\right), \quad (34\cdot22)$$

$$(\mathcal{L}_z)_{\text{op}} = \frac{h}{2\pi i}\,\frac{\partial}{\partial\varphi}. \quad (34\cdot23)$$

Equations (34·21) to (34·23) can be derived by direct transformation or by the following less tedious considerations [*cf.* G. Breit, "Separation of Angles in the Two-electron Problem," *Phys. Rev.* **35**, 569–578 (1930)]. Let F denote an arbitrary differentiable function of the generalized coordinates $q_1, \cdots, q_s$. A class of linear differential operators applicable to F is defined by the equation

$$WF(q_1, \cdots, q_s) \equiv \sum_k v_k\frac{\partial F}{\partial q_k} \equiv \left[\frac{d}{dt}F(q_1 + v_1 t, \cdots, q_s + v_s t)\right]_{t=0}. \quad (a)$$

The operator W depends on the s independent quantities $v_1, v_2, \cdots, v_s$ which may be arbitrary functions of the coordinates $q_1, \cdots, q_s$. The v's can be thought of as velocity components of a fluid motion in the configuration space $q_1, \cdots, q_s$.

The *velocity operators* W have the important property that

$$W(v_1, \cdots, v_s) + W(v_1', \cdots, v_s') = W(v_1 + v_1', \cdots, v_s + v_s'). \quad (b)$$

Furthermore, if we wish to express F and W in terms of a second set of generalized coordinates $Q_1, \cdots, Q_s$, we have only to write

$$\left.\begin{aligned}F(q_1, \cdots, q_s) &= G(Q_1, \cdots, Q_s), \qquad WG = \sum_k V_k\frac{\partial G}{\partial Q_k}, \\ V_k &= \left[\frac{dQ_k}{dt}\right]_{t=0} = \sum_j v_j\frac{\partial Q_k}{\partial q_j}.\end{aligned}\right\} \quad (c)$$

The linear-momentum operators are obviously complex multiples of special velocity operators corresponding to uniform translations along the x, y, and z axes, respectively. Similarly, $\frac{2\pi i}{h}(\mathcal{L}_x)_{\text{op}}, \frac{2\pi i}{h}(\mathcal{L}_y)_{\text{op}}, \frac{2\pi i}{h}(\mathcal{L}_z)_{\text{op}}$ are special cases in which the velocity system is that of a uniform rotation of unit angular velocity about the x, y, and z axes, respectively. Finally the operators $\partial/\partial\varphi$, $\partial/\partial\psi$, $\partial/\partial\theta$ in the coordinate system φ, ψ, θ, z_1', x_2', z_2', $\cdots$, z_s' are velocity operators for rigid rotations of the system of particles about the z axis, the z' axis, and the line of nodes, respectively.

We can now make use of a fundamental theorem regarding the compounding of angular velocities in order to express $(\mathcal{L}_x)_{\text{op}}$, $(\mathcal{L}_y)_{\text{op}}$, $(\mathcal{L}_z)_{\text{op}}$ as linear combinations of the operators $\partial/\partial\varphi$, $\partial/\partial\psi$, $\partial/\partial\theta$ and *vice versa*. Let $\overrightarrow{\Omega}_1$ and $\overrightarrow{\Omega}_2$ denote two angular velocities with the resultant $\overrightarrow{R}$. Let $\overrightarrow{V}_1$ and $\overrightarrow{V}_2$ denote the vector linear velocities correlated with $\overrightarrow{\Omega}_1$ and $\overrightarrow{\Omega}_2$. Then, according to this theorem, the linear velocity system V for the equivalent single angular velocity $\overrightarrow{R}$ is

$$\overrightarrow{V} = \overrightarrow{V}_1 + \overrightarrow{V}_2.$$

It follows from Eq. (*b*) that the velocity operator for the angular velocity $\overrightarrow{R}$ is the sum of the velocity operators for $\overrightarrow{\Omega}_1$ and $\overrightarrow{\Omega}_2$. Thus the velocity operators corresponding

to different rigid rotations of the system of particles can be combined just like the corresponding angular-velocity vectors.

Let the second position of the η axis in passing from the x,y,z orientation to the x',y',z' orientation be designated by η_2. This has the direction of the angular velocity associated with $\partial/\partial\theta$. A unit vector in this direction is equal to the sum of vectors of length $\cos(x,\eta_2)$, $\cos(y,\eta_2)$, $\cos(z,\eta_2)$ in the directions of the x, y, and z axes, respectively. But the operator corresponding to an angular velocity of magnitude $\cos(x,\eta_2)$ in the direction of the x axis is $(2\pi i/h)\cos(x,\eta_2)(\mathcal{L}_x)_{\text{op}}$. Hence the operator equation corresponding to the resolution of a vector in the direction of the η_2 axis into vector components along the coordinate axes is

$$\frac{h}{2\pi i}\frac{\partial}{\partial\theta} = \cos(x,\eta_2)(\mathcal{L}_x)_{\text{op}} + \cos(y,\eta_2)(\mathcal{L}_y)_{\text{op}} + \cos(z,\eta_2)(\mathcal{L}_z)_{\text{op}}$$

$$= -\sin\varphi\,(\mathcal{L}_x)_{\text{op}} + \cos\varphi\,(\mathcal{L}_y)_{\text{op}}.$$

A similar resolution of the angular-velocity systems corresponding to the operators $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial\varphi}$ and $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial\psi}$ yields

$$\frac{h}{2\pi i}\frac{\partial}{\partial\varphi} = (\mathcal{L}_z)_{\text{op}}$$

$$\frac{h}{2\pi i}\frac{\partial}{\partial\psi} = \cos\varphi\sin\theta\,(\mathcal{L}_x)_{\text{op}} + \sin\varphi\sin\theta\,(\mathcal{L}_y)_{\text{op}} + \cos\theta\,(\mathcal{L}_z)_{\text{op}}.$$

Equations (34·21), (34·22), (34·23) are readily obtained by inverting the above equations.

Squaring and adding these operators we obtain the following characteristic equation for $\mathcal{L}^2$ to be solved subject to the continuity condition and the requirement of quadratic integrability:

$$\mathcal{L}^2 U = -\frac{h^2}{4\pi^2}\left\{\left(1 + \frac{\cos^2\theta}{\sin^2\theta}\right)\frac{\partial^2 U}{\partial\varphi^2} + \frac{\partial^2 U}{\partial\theta^2} + \frac{1}{\sin^2\theta}\frac{\partial^2 U}{\partial\psi^2} - \frac{2\cos\theta}{\sin^2\theta}\frac{\partial^2 U}{\partial\varphi\partial\psi} \right.$$
$$\left. + \frac{\cos\theta}{\sin\theta}\frac{\partial U}{\partial\theta}\right\} = (\mathcal{L}^2)'U. \quad (34\cdot24)$$

Here we indicate the wave function by the symbol U to avoid confusion with the Eulerian angle ψ. As the operator $\mathcal{L}^2$ does not contain the relative coordinates z_1', x_2', $\cdots$, z_f' the eigenfunctions of $\mathcal{L}^2$ are determined by Eq. (34·24) only to a factor which is an arbitrary function of the primed coordinates. To solve Eq. (34·24) we set

$$U = e^{i(M\varphi+N\psi)}P(\theta)Q(z_1', \cdots z_f') \qquad M, N = 0, \pm 1, \pm 2, \cdots \quad (34\cdot25)$$

Substitution yields

$$\frac{d^2P}{d\theta^2} + \frac{\cos\theta}{\sin\theta}\frac{dP}{d\theta} + \left[\frac{4\pi^2(\mathcal{L}^2)'}{h^2} - N^2 - \frac{(N\cos\theta - M)^2}{\sin^2\theta}\right]P = 0. \quad (34\cdot26)$$

This equation is a generalization of Eq. (28·8), to which it reduces if we set N equal to zero. It can be solved by the polynomial methods used in treating (28·8) and has been so treated by Sommerfeld.[1] Equation (34·24) is in fact essentially the wave equation for the free symmetric

[1] SOMMERFELD, *Atombau und Spektrallinien, Wellenmechanischer Ergänzungsband,* Kap. I, §110, Braunschweig, 1929.

top.[1]   The eigenvalues of $\mathcal{L}^2$ are the same as in the one-electron problem, viz., $L(L + 1)h^2/4\pi^2$, where $L = 0, 1, 2 \cdots$.   The minimum value of the azimuthal or angular-momentum quantum number L for any given values of M and N is the larger of the two numbers $|M|$ and $|N|$.   It follows that for a given values of L, M can take on all values between $-L$ and $+L$.   The eigenvalue $(\mathcal{L}^2)' = L(L + 1)h^2/4\pi^2$ has $(2L + 1)$-fold degeneracy.

Changing the independent variable from θ to the quantity t defined by

$$t = \tfrac{1}{2}(1 - \cos \theta)$$

permits us to express the eigenfunctions $P_{MN}^L(\theta)$ in terms of Jacobi polynomials.   Introducing the conventional notation for these polynomials and the definitions

$$d = |M - N|, \qquad s = |M + N|, \qquad 2p = 2L - (d + s),$$

we have

$$P_{MN}^L(\theta) = t^{\frac{d}{2}}(1 - t)^{\frac{s}{2}}G_p(1 + d + s, 1 + d, t). \tag{34·27}$$

A summary of the more important properties of the Jacobi polynomials is given in Appendix J.

Since the eigenvalues of $\mathcal{L}^2$ range from zero to infinity, it is not difficult to show that the series of eigenfunctions for all possible values of L consistent with any given values of M and N forms a complete system. As the exponential functions also form complete systems, it is possible to expand an arbitrary continuous function of θ, φ, ψ, spread out over the region $0 \leqslant \varphi \leqslant 2\pi$, $0 \leqslant \theta \leqslant \pi$, $0 \leqslant \psi \leqslant 2\pi$ into a mean-square convergent series of the form

$$f(\varphi, \theta, \psi) = \sum_{L > \left\{ \begin{smallmatrix} |M| \\ |N| \end{smallmatrix} \right.}^{+\infty} \sum_{N = -\infty}^{+\infty} \sum_{M = -\infty}^{+\infty} c_{LMN}\, e^{i(M\varphi + N\psi)} P_{MN}^L(\theta). \tag{34·28}$$

35. THE ENERGY OPERATORS

35a. Calculation of Probabilities and Mean Values of Energy.—The expansion (30·32) resolves an arbitrary normal wave function $\Psi(x,t)$ into a series-integral combination of eigenfunctions of the energy operator H.   Each term of this expansion is a harmonic function of the time t with the frequency $\nu_n = E_n/h$.   In optics the square of the wave amplitude, which we may call the intensity of the wave system, measures the energy density.   The integrated intensity of each monochromatic component gives the total energy of the corresponding frequency.   If we divide this integrated intensity by the corresponding photon energy

[1] *Cf., e.g.,* CONDON and MORSE, *Q.M.*, pp. 74–77.

$h\nu$, we obtain the total number of photons of the given frequency in the assemblage described by the wave function. In our present theory the intensity of the waves, $|\Psi|^2$, measures probability density directly, and by analogy we should expect that if we resolve Ψ into monochromatic components and compute the integrated intensity of any of them, the result will measure the total probability of the corresponding frequency and energy. If we have an assemblage of systems in a definite state described by a normalized wave function Ψ and multiply the integrated intensity of the waves of a given frequency by the total number of systems in the assemblage, we obtain the most probable number of systems having the corresponding energy. If the number so calculated is very large, we can identify the most probable number with the actual number found by measurement in any special case without serious likelihood of appreciable error. This method of computing the probability of an energy value is generalized in Sec. 36 to apply to the probabilities of other "dynamical variables" and is subject to further scrutiny in Sec. 41.

In order to reduce the above rule to a formula, we assume that the normalized wave function Ψ is subjected to an expansion of the form of (32·28). By performing the expansion appropriate to the time $t = 0$ and multiplying each term by the appropriate time factor $e^{-\frac{2\pi i Et}{h}}$, we obtain an expansion suitable for an arbitrary time t [cf. Eqs. (30·36) and (30·42)]. Thus

$$\Psi(x,t) = \sum_k \sum_n{}' c_{kn}\psi_{kr}(x)e^{-\frac{2\pi i E_k t}{h}} + \sum_n \int c_n(E)\psi_n(E,x)e^{-\frac{2\pi i Et}{h}}\, dE. \quad (35\cdot1)$$

The partial wave function for the discrete energy E_k is $\sum_n{}' c_{kn}\psi_{kn}e^{-\frac{2\pi i E_k t}{h}}$.

By (32·27) its integrated intensity is $\sum_n{}' c_{kn}c_{kn}{}^*$. If Ψ_k denotes the normalized value of the above partial wave function, while N_k and N give, respectively, the number of systems of energy E_k and the total number, we have

$$\frac{N_k}{N} = \sum_n c_{kn}c_{kn}{}^* = \sum_n |(\Psi,\psi_{kn})|^2 = |(\Psi,\Psi_k)|^2. \quad (35\cdot2)$$

Similarly the fraction of systems in the infinitesimal interval dE of the continuous range of energy values is $\sum_n{}' c_n(E)c_n(E){}^*dE$, assuming appropriate normalization of the corresponding wave functions. Thus the

set of discrete coefficients c_{kn} together with the functions $c_n(E)$ play the part of probability amplitude for energy.[1]

The above procedure for calculating the probability of different energies is obviously similar in form to the scheme used for computing the probabilities of different linear momenta. In fact it can be derived from the latter in the special case of free particles, where the energy is a function of the momentum only, independent of position. In this case the analysis of Ψ into plane waves is also an analysis into components of definite frequency and energy. This is evident from (15·8). The number of particles of energy intermediate between E' and $E' + dE'$ is equal to the total number of particles with momenta p_x, p_y, p_z which give points in the spherical shell

$$E' \leqslant \frac{(p_x^2 + p_y^2 + p_z^2)}{2\mu} \leqslant E' + dE'$$

in momentum space. By the analysis of Sec. 15, this number is equal to $\int_{shell} |\Phi|^2 dp_x dp_y dp_z$, which is nothing more or less than the integrated intensity of the residual wave function in x,y,z space when all harmonic components have been removed except those which satisfy the inequality $E' \leqslant E \leqslant E' + dE'$.

It follows from (35·1) and Eqs. (32·29) and (32·30) that

$$H\Psi = -\frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} = \sum_k \sum_n c_{kn} E_k \psi_{kn} e^{-\frac{2\pi i E_k t}{h}}$$

$$+ \sum_n \int c_n(E) E \psi_n(E,x) e^{-\frac{2\pi i E t}{h}} dE. \qquad (35\cdot3)$$

Hence

$$\bar{E} = \int_\infty \Psi^* \left(-\frac{h}{2\pi i}\frac{\partial}{\partial t}\right)\Psi d\tau = \int_\infty \Psi^* H\Psi d\tau. \qquad (35\cdot4)$$

[1] It is easy to jump at the conclusion that $\sum_n |c_{kn}\psi_{kn}|^2 d\tau$ is the probability that an arbitrary system will simultaneously have the energy E_k and a configuration belonging to the volume element $d\tau$—a mistake made by Kemble and Hill, *Rev. Mod. Phys*, **2**, 6 (1930). It is true that it does give the fraction of the original number of systems remaining, if one first eliminates all energy values but E_k, and later eliminates all coordinate values outside $d\tau$, but the observation of the energy alters the position and *vice versa*, so that one cannot use the energy together with the space coordinates to form a single coordinate system.

As an example of a proper coordinate system involving the energy we may mention the three quantities E, $\mathcal{L}^2$, $\mathcal{L}_z$ in the case of the two-particle problem. The expansion (30·35) resolves the wave function in this case into a linear combination of functions of all three quantities. The system of coefficients forms the probability amplitude for the coordinate system E, $\mathcal{L}^2$, $\mathcal{L}_z$.

Thus the equivalent operators $-\dfrac{h}{2\pi i}\dfrac{\partial}{\partial t}$ and H bear the same relation to

the numerical values of E as $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial x}$ bears to the numerical values of p_x.

(As stated on p. 26, classically $-E$ can be regarded as the momentum conjugate to the time t when the latter is treated as a coordinate.)   In the case of a sharply defined wave packet the argument of footnote 1, p. 229, can be used to prove that the mean energy derived from (35·4) is equal to the classical energy $H(p,q)$.

*35b. Transformation of Hamiltonian Operator.—The disclosure of the intimate relation between the Hamiltonian function of classical theory in Cartesian coordinates and the fundamental operator of the Schrödinger equation raises the question of the relation between these expressions in other coordinate systems.[1]   It is simpler to transform the classical Hamiltonian from one coordinate system to another than to make the corresponding direct transformation of the Hamiltonian operator.   Hence a rigorous method for passing from the classical function

$H(p,q)$ in generalized coordinates to the appropriate operator $H\!\left(\dfrac{\partial}{\partial q_k},q_k\right)$

would be useful.

A natural conjecture regarding the answer to this question would be

that we simply employ the substitution $p_k \to \dfrac{h}{2\pi i}\dfrac{\partial}{\partial q_k}$ again.   This rule is

sufficient to guarantee that H shall give the classical energy in the case of a well-defined wave packet, but does not always give the same result as direct transformation of the primary operator from the Cartesian system to the system of coordinates desired.   Consider, for example, the kinetic term $p_r{}^2/2\mu$ of the classical Hamiltonian function for a single particle in spherical coordinates.   So long as p_r is a number, the factor

$p_r{}^2$ can be replaced by either $\dfrac{1}{f(r)}p_r f(r)p_r$ or $\dfrac{1}{f(r)}p_r{}^2 f(r)$ with $f(r)$ arbitrary,

but when we convert p_r into an operator these various forms are no longer equivalent.   Direct transformation of the Cartesian operator yields the radial term [*cf.* Eq. (28·3)]

$$-\frac{h^2}{8\pi^2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}\!\left(r^2\frac{\partial}{\partial r}\right),$$

which is obtainable by our simple substitution rule, only if we start from the special classical form $r^{-2}p_r r^2 p_r$.   We are thus driven back on a direct transformation of the Hamiltonian operator, or, what comes to the same thing, of the Laplacian operator, from Cartesian to generalized coordinates.

[1] *Cf.* E. Schrödinger, *Ann. d. Physik* (4) **79**, 747–748 (1926); Courant-Hilbert, *M.M.P.*, §8, pp. 192–195; B. Podolsky, *Phys. Rev.* **32**, 812 (1928).

This transformation is sometimes facilitated by the use of formulas derivable from the variation principle (or otherwise). Let us first introduce a weighted set of Cartesian coordinates X_1, X_2, $\cdots$, X_λ defined by the equations

$$X_1 = \mu_1^{1/2}x_1, \qquad X_2 = \mu_1^{1/2}y_1, \qquad X_3 = \mu_1^{1/2}z_1, \qquad X_4 = \mu_2^{1/2}x_2, \quad \cdots .$$

These transform the operator $\displaystyle\sum_k \frac{1}{\mu_k}\nabla_k^2$ into the λ-dimensional Laplacian

$$\nabla^2 = \sum_{k=1}^{\lambda} \frac{\partial^2}{\partial X_k^2}.$$

Let the final set of generalized coordinates be designated by the symbols q_1, q_2, $\cdots$,q_λ. Let the symbols g_{ik} and g^{ik} denote the quantities

$$g_{ik} = \sum_{\tau=1}^{\lambda} \frac{\partial X_\tau}{\partial q_i}\frac{\partial X_\tau}{\partial q_k}, \qquad g^{ik} = \sum_{\tau=1}^{\lambda} \frac{\partial q_i}{\partial X_\tau}\frac{\partial q_k}{\partial X_\tau}, \tag{35·5}$$

respectively. Let g denote the determinant

$$g = \begin{vmatrix} g_{11} & g_{12} & \cdots & g_{1\lambda} \\ g_{21} & g_{22} & \cdots & g_{2\lambda} \\ \cdots & \cdots & \cdots & \cdots \\ g_{\lambda 1} & g_{\lambda 2} & \cdots & g_{\lambda\lambda} \end{vmatrix}.$$

($g^{1/2}$ is the functional determinant, or Jacobian, of the coordinates X_1, $\cdots$ X_λ with respect to the system q_1, $\cdots$ q_λ). Then the λ-dimensional Laplacian of ψ in the q system is

$$\nabla^2\psi = \frac{1}{\sqrt{g}}\sum_i \frac{\partial}{\partial q_i}\left(\sqrt{g}\sum_k g^{ik}\frac{\partial\psi}{\partial q_k}\right), \tag{35·6}$$

and the Schrödinger equation in the q system is

$$H\psi = -\frac{h^2}{8\pi^2}\frac{1}{g^{1/2}}\sum_i \frac{\partial}{\partial q_i}\left(g^{1/2}\sum_k g^{ik}\frac{\partial\psi}{\partial q_k}\right) + V\psi = E\psi. \tag{35·7}$$

It is convenient to designate the wave function of Eq. (35·7) by the symbol ψ_X to indicate the fact that its normalization condition is by hypothesis

$$\int_\infty \psi_X\psi_X^* dX_1 \cdots dX_\lambda = 1. \tag{35·8}$$

ψ_X can be expressed either in terms of the original Cartesian coordinates, or in terms of the q's, but, whereas $|\psi_X|^2 dX_1 \cdots dX_\lambda$ gives the proba-

bility of the element $dX_1 dX_2 \cdots dX_\lambda$ of the X coordinate space, we must multiply $|\psi_x|^2$ by the Jacobian $g^{1/2}$ in order to obtain the probability density in q space. Thus the normalization integral for ψ_x in q space has the form $\int \psi_x \psi_x^* g^{1/2} dq_1 \cdots dq_\lambda$. This extra factor $g^{1/2}$ is sometimes called the *density factor* for ψ_x in q space.

Let us now introduce a new wave function ψ_q defined by the formula

$$\psi_q = g^{1/4} \psi_x. \tag{35.9}$$

The normalization condition for ψ_q is

$$\int \psi_q \psi_q^* dq_1 \cdots dq_\lambda = \int \psi_q \psi_q^* g^{-1/2} dX_1 \cdots dX_\lambda = 1. \tag{35.10}$$

Thus ψ_q has the density factor unity in q space, whereas ψ_x has the density factor unity in X space. The substitution of $g^{-1/4}\psi_q$ for ψ_x in Eq. (35.7) gives the following differential equation for ψ_q:

$$H_q \psi_q = -\frac{h^2}{8\pi^2} g^{-1/4} \sum_i \frac{\partial}{\partial q_i} \left(g^{1/2} \sum_k g^{ik} \frac{\partial}{\partial q_k} g^{-1/4} \psi_q \right) + V\psi_q$$

$$= -\frac{h^2}{8\pi^2} \sum_i \sum_k \left\{ g^{ik} \frac{\partial^2 \psi_q}{\partial q_i \partial q_k} + \frac{\partial g^{ik}}{\partial q_i} \frac{\partial \psi_q}{\partial q_k} - \frac{1}{4} g^{-1/4} \frac{\partial}{\partial q_i} \left(g^{-3/4} g^{ik} \frac{\partial g}{\partial q_k} \right) \psi_q \right\}$$

$$+ V\psi_q = E\psi_q. \tag{35.11}$$

Example.—As an illustration of the application of the above formulas let us consider the transformation of the wave equation for the three-particle problem (the coordinates of the center of gravity eliminated) from the Cartesian coordinates of Eq. (32.17) to the spherical coordinates of Eq. (32.18). We reletter the coordinates as follows:

$$
\begin{array}{llll}
X_1 = \xi_1, & X_4 = \xi_2, & q_1 = R, & q_4 = \theta_2, \\
X_2 = \eta_1, & X_5 = \eta_2, & q_2 = \alpha, & q_5 = \varphi_1, \\
X_3 = \zeta_1, & X_6 = \zeta_2, & q_3 = \theta_1, & q_6 = \varphi_2.
\end{array}
$$

The transformation equations are then

$$
\begin{array}{ll}
X_1 = q_1 \cos q_2 \sin q_3 \cos q_5, & X_4 = q_1 \sin q_2 \sin q_4 \cos q_6, \\
X_2 = q_1 \cos q_2 \sin q_3 \sin q_5, & X_5 = q_1 \sin q_2 \sin q_4 \sin q_6, \\
X_3 = q_1 \cos q_2 \cos q_3, & X_6 = q_1 \sin q_2 \cos q_4.
\end{array}
$$

The q system of coordinates is an orthogonal one, *i.e.*, one for which the determinant g takes a diagonal form owing to the fact that all g_{ik}'s vanish for which $i \neq k$. The general formulas,

$$\sum_i g_{ik} g^{il} = \begin{cases} 0 \text{ for } k \neq 1, \\ 1 \text{ for } k = 1, \end{cases} \qquad \bullet \qquad \tag{35.12}$$

then show that $g^{ik} = 0$ if $i \neq k$ and that $g^{kk} = 1/g_{kk}$. From Eqs. (35·5).
we obtain

$$
\begin{aligned}
g_{11} &= 1, & g_{44} &= q_1{}^2 \sin^2 q_2, \\
g_{22} &= q_1{}^2, & g_{55} &= q_1{}^2 \cos^2 q_2 \sin^2 q_3, \\
g_{33} &= q_1{}^2 \cos^2 q_2, & g_{66} &= q_1{}^2 \sin^2 q_2 \sin^2 q_4, \\
g^{1/2} &= q_1{}^5 \sin^2 q_2 \cos^2 q_2 \sin q_3 \sin q_4.
\end{aligned}
$$

Equation (35·6) yields

$$
\nabla^2 \psi = \frac{1}{\sqrt{g}} \sum_k \frac{\partial}{\partial q_k}\left(\frac{\sqrt{g}}{g_{kk}} \frac{\partial \psi}{\partial q_k}\right)
$$

$$
= \frac{1}{q_1{}^5} \frac{\partial}{\partial q_1}\left(q_1{}^5 \frac{\partial \psi}{\partial q_1}\right) + \frac{1}{q_1{}^2 \sin^2 q_2 \cos^2 q_2} \frac{\partial}{\partial q_2}\left(\sin^2 q_2 \cos^2 q_2 \frac{\partial \psi}{\partial q_2}\right)
$$

$$
+ \frac{1}{q_1{}^2 \cos^2 q_2 \sin q_3} \frac{\partial}{\partial q_3}\left(\sin q_3 \frac{\partial \psi}{\partial q_3}\right) + \frac{1}{q_1{}^2 \sin^2 q_2 \sin q_4} \frac{\partial}{\partial q_4}\left(\sin q_4 \frac{\partial \psi}{\partial q_4}\right)
$$

$$
+ \frac{1}{q_1{}^2 \cos^2 q_2 \sin^2 q_3} \frac{\partial^2 \psi}{\partial q_5{}^2} + \frac{1}{q_1{}^2 \sin^2 q_2 \sin^2 q_4} \frac{\partial^2 \psi}{\partial q_6{}^2}. \quad (35\cdot13)
$$

The resulting wave equation, after reverting to the original symbols, is
(32·18), if D is given the explicit form,

$$
D = \frac{1}{\sin^2 \alpha \cos^2 \alpha} \frac{\partial}{\partial \alpha}\left(\sin^2 \alpha \cos^2 \alpha \frac{\partial}{\partial \alpha}\right) + \frac{1}{\cos^2 \alpha \sin \theta_1} \frac{\partial}{\partial \theta_1}\left(\sin \theta_1 \frac{\partial}{\partial \theta_1}\right)
$$

$$
+ \frac{1}{\cos^2 \alpha \sin^2 \theta_1} \frac{\partial^2}{\partial \varphi_1{}^2} + \frac{1}{\sin^2 \alpha \sin \theta_2} \frac{\partial}{\partial \theta_2}\left(\sin \theta_2 \frac{\partial}{\partial \theta_2}\right) + \frac{1}{\sin^2 \alpha \sin^2 \theta_2} \frac{\partial^2}{\partial \varphi_2{}^2}.
$$

$$(35\cdot14)$$

For ψ_q, Eq. (35·11) gives

$$
H\psi_q = -\frac{h^2}{8\pi^2 \mu} g^{-1/4} \sum_k \frac{\partial}{\partial q_k}\left(\frac{g^{1/2}}{g_{kk}} \frac{\partial}{\partial q_k} g^{-1/4} \psi_q\right) + V(q)\psi_q
$$

$$
= -\frac{h^2}{8\pi^2 \mu}\left\{\frac{\partial^2}{\partial R^2} + \frac{1}{R^2}\left[\frac{\partial^2}{\partial \alpha^2} + \frac{1}{\cos^2 \alpha}\left(\frac{\partial^2}{\partial \theta_1{}^2} + \frac{1}{\sin^2 \theta_1} \frac{\partial^2}{\partial \varphi_1{}^2} + \frac{1 + \csc^2 \theta_1}{4}\right)\right.\right.
$$

$$
\left.\left. + \frac{1}{\sin^2 \alpha}\left(\frac{\partial^2}{\partial \theta_2{}^2} + \frac{1}{\sin^2 \theta_2} \frac{\partial^2}{\partial \varphi_2{}^2} + \frac{1 + \csc^2 \theta_2}{4}\right) + \frac{1}{16}\right]\right\}\psi_q + V(q)\psi_q = E\psi_q.
$$

$$(35\cdot15)$$

It will be noted that all the cross-derivatives are eliminated by the
transition from $\psi_x(q)$ to $\psi_q(q)$. This is true whenever the q system is an
orthogonal one as may be seen from Eq. (35·11).

36. DYNAMICAL VARIABLES IN GENERAL

36a. Remarks on the Value of the General Theory.—The foregoing
discussion of the operators for linear momentum, angular momentum,
and energy cries out for generalization. One would like to know whether

every measurable physical quantity α depending on the state of a dynamical system is to be associated with a corresponding operator $(\alpha)_{op}$ and, if so, what properties are to be assigned to the class of operators representing measurable physical quantities.

It must be frankly admitted at the outset that an examination of the efforts which have been made to set up a general theory answering the above and related questions suggests at times the possibility that the game is not worth the candle. One of the chief fruits of this quest is the Dirac-Jordan transformation theory, which gratifies our natural desire for unity and completeness and has considerable practical value, but like Dirac's later formalism[1] is of too heuristic a character to be ultimately satisfying. On the other hand the mathematically rigorous and elegant method of von Neumann involves a technique at once too delicate and too cumbersome for the practical purposes of the average physicist.

One's doubts about these developments, with their emphasis on dynamical variables defined by arbitrary Hermitian operators, are emphasized by an examination of the dynamical variables actually measured for atomic systems. As a matter of fact, position, or configuration, linear momentum, energy, and a single arbitrary component of magnetic moment are the only independent dynamical variables whose measurement can be carried out in principle with arbitrary precision for an individual atomic system. As the magnetic moment is determined by the angular momentum and energy, one can say that the only variables which can be defined unambiguously in terms of operations we can approximate in the laboratory are those already considered, together with other quantities which are functions of them. Hence it becomes evident that all of the practical results of nonrelativistic quantum theory might be derived with the aid of the explicit discussion of the small group of dynamical variables mentioned and their operators, without a general theory.

Nevertheless, it is to be remembered that the development of scientific theory is always conditioned by artistic considerations of simplicity and symmetry, and by the urge for unity and completeness. Hence a proper exposition of modern physical theory ought to include more than a bare analysis of the relation between the essential postulates and the experiments which have been performed or which may soon be performed. It must include an examination of efforts to reduce the theory to a compact, unified, and general scheme, even when these efforts are only partially successful. In fact such an examination is not only desirable from a philosophical standpoint but is a practical necessity for the student to whom much of the literature of the subject would otherwise remain a closed book. Moreover, the attempts at generalization form an essential background for any endeavor to develop the theory to include

---

[1] *I.e.,* the formalism of his book on quantum mechanics.

a wider range of experimental facts. The reader is therefore invited to join the author in an attempt to generalize and unify the results already obtained.

36b. Possibility of Defining Physical Quantities by Operators.—In view of the above remarks concerning the dynamical variables which are actually measurable for atomic systems—or for any system, provided the accuracy is sufficient to take the measurement out of the domain of classical physics into that of quantum physics—we shall *assume* that the answer to the first of the two questions raised at the beginning of the last section is affirmative. To be specific we postulate that every dynamical variable α defined operationally by an exact scheme of measurement can be correlated with a corresponding linear operator $(\alpha)_{op}$. Postponing to Chap. IX an examination of the relation between the scheme of measurement and the properties of the operator in question, we turn our attention next to the question of the common characteristics of the class of operators $(\alpha)_{op}$ associated with actually, or potentially, measurable physical quantities.

These common characteristics can be inferred by induction from the special cases of positional coordinates, linear momentum, angular momentum, and energy previously examined. In every case the operator $(\alpha)_{op}$ is linear and Hermitian[1] with respect to a class of functions which includes class D and which in the language of von Neumann (*cf.* Sec. 32c) is "everywhere dense" in the Hilbert space of all quadratically integrable functions. In all cases *except those of the positional coordinates* there exists a complete discrete-continuous set of solutions of the equation

$$(\alpha)_{op}\varphi(x) = \alpha'\varphi(x) \tag{36·1}$$

in terms of which an arbitrary quadratically integrable function of the coordinates, say $\psi(x)$, is mean-square expansible. The eigenvalues α' associated with the elements of this set have been identified with the values of the variable α which may result from an experimental measurement. They are all real, due to the Hermitian character of $(\alpha)_{op}$.[2] The expectation value $\bar{\alpha}$ for measurements of systems belonging to an assemblage in a state described by the wave function[3] $\psi(x)$ is given in all cases by the formula $\int\psi^*(\alpha)_{op}\psi d\tau = ((\alpha)_{op}\psi,\psi)$. To determine the proba-

---

[1] For certain purposes it is convenient to include with the Hermitian dynamical variable operators whose eigenvalues are real, other non-Hermitian operators with complex eigenvalues. An appropriate generalization of the definition of a quantum-mechanical dynamical variable is given in Sec. 36j. For the present we restrict our attention to the real, or Hermitian, case.

[2] The case where α' is a discrete eigenvalue is covered by the argument of p. 206. The case in which α belongs to a continuous spectrum will be dealt with on p. 275.

[3] In this section the dependency of the wave functions on the time is of no importance. Hence we use the symbol ψ for an arbitrary wave function, or instantaneous Ψ.

bility of any discrete eigenvalue, say α_n, one expands $\psi(x)$ into a discrete-continuous linear combination of eigenfunctions and evaluates the norm, $N[c_n\psi_n] = (c_n\psi_n, c_n\psi_n)$, of the term (or sum of terms) in the expansion which belongs to α_n. This norm is the probability in question. In the case of continuous-spectrum eigenvalues one evaluates the probability of a small range of eigenvalues $d\alpha'$ by a similar process. Thus $(\alpha)_{op}$ determines the eigenvalues and unites with the state function $\psi(x)$ to fix their probabilities.

In the case of the positional coordinates the allowed values and the probabilities of different ranges are fixed by our initial hypothesis regarding the physical interpretation of $|\psi|^2$ and the assumption that ψ is a continuous function of the Cartesian coordinates. Hence it was not necessary to use an operator to find the eigenvalues and their probabilities. However, the formula (33·1), or its instantaneous equivalent

$$\bar{q} = (q\psi, \psi),$$

suggests the propriety of identifying $(q)_{op}$ with the multiplication operator $[q \times]$, and implies that this operator, along with ψ, must determine the probabilities of different ranges of q values. The properties of the class of operators for the positional coordinates thus defined turn out to be sufficiently different from those of the other Hermitian operators studied, so that, strictly speaking, we cannot use the above described eigenvalue-eigenfunction method to predict the results of positional measurements made on systems in a state described by the known wave function $\psi(x)$. Nevertheless, this difficulty can be overcome by a proper reformulation of the eigenvalue-eigenfunction problem, to which we shall later return. With the aid of such a reformulation it can be shown that the positional coordinates are no exception to the general rule that the operator associated with a dynamical variable determines its spectrum and, when used in conjunction with the wave function $\psi(x)$, can be employed to work out the appropriate distribution function giving the probabilities of different measured values.

It will be observed that, inasmuch as the operator $(\alpha)_{op}$ in any of these cases determines the result of measuring α for a large number of systems in any common definite state, it must implicitly determine the experimental procedure appropriate to measuring α. Hence we may say that a physical quantity α can be defined by the corresponding operator $(\alpha)_{op}$.

Now it is customary to define the method of measuring physical quantities without defining these quantities themselves. In fact we have no satisfactory reason for ascribing objective existence to physical quantities as distinguished from the numbers obtained when we make the measurements which we correlate with them. As indicated in Sec. 19d, there is no real reason for supposing that a particle belonging to an assemblage described by a wave function $\psi(x)$ has at every moment a

definite, but unknown, position which may be revealed by a measurement of the right kind, or a definite momentum which can be revealed by a different measurement. On the contrary, we get into a maze of contradictions as soon as we inject into quantum mechanics such concepts carried over from the language and philosophy of our scientific ancestors. As no scheme of operations can determine experimentally whether physical quantities such as position and momentum exist and have unique values when they are not at the moment under observation, nor whether the number obtained by a measurement describes some objective property of the thing measured, a strict adherence to the operational point of view requires that we eliminate such concepts from our theories. From the standpoint of classical physics such a rejection would perhaps be a bit of philosophical purism flying so unnecessarily in the face of common sense that few would care to adopt it. On the other hand this rejection is a demonstrable logical necessity for quantum mechanics.[1]

We bring up this philosophical question here because of its relation to our notation. It would evidently be philosophically more exact if we spoke of "making measurements" of this, that, or the other type instead of saying that we measure this, that, or the other "physical quantity." Rather than make such a radical and awkward change in our phraseology, however, we can continue to use the old language, reinterpreting the terms "physical quantity" and "dynamical variable," and allowing them to stand for the corresponding operator which fixes the nature of the measurement under consideration. We here adopt this latter procedure as a matter of convenience, interpreting the phrase "measurement of the z component of angular momentum," for example, as equivalent to "making eigenvalue observations for the operator $(\mathcal{L}_z)_{op}$." In so doing we refuse to recognize any content to the symbol α which does not flow out of the definition of the operator $(\alpha)_{op}$ and of the boundary and continuity conditions which fix its eigenvalues and eigenfunctions. (The boundary and continuity conditions are themselves ultimately, as we shall see, implicit in the complete definition of the operator.) Hence we have no real need for the symbol α as distinguished from $(\alpha)_{op}$ and can logically discard the complication of the latter symbol. *Henceforth we shall accordingly drop our original operator notation and substitute for the operator symbol $(\alpha)_{op}$ the symbol α itself.*

The distinction between operators and their eigenvalues will not be entirely uniform. In some cases we shall use Greek letters for operators and Roman for their eigenvalues, but usually the eigenvalues will be

---

[1] This fact has been clearly brought out in the discussion of a recent paper on quantum mechanics and physical reality by Einstein, Podolsky, and Rosen [*Phys. Rev.* **47**, 777 (1935)]. See especially the paper by W. H. Furry, *Phys. Rev.* **49**, 393 (1936).

indicated by the same symbol as that used for the operator, but with a prime or a subscript as a distinguishing mark.

36c. The Transformation of Probability Amplitudes and Dynamical Variables.—It is necessary to add that the concrete mathematical form of the operator corresponding to a given type of physical measurement must vary with the nature of the scheme of probability amplitudes for which it is to be used. Thus we have already seen that the operators H, $\mathcal{L}^2$, $\mathcal{L}_z$ take on different forms for different coordinate systems. In fact every transformation of the wave function used to describe an assemblage of dynamical systems involves a corresponding transformation of the operator for each "dynamical variable."

If we know how to transform probability amplitudes from one coordinate system to another, it is a simple matter to derive corresponding transformation formulas for the operators associated with a given dynamical variable. Let ψ_x denote a wave function or probability amplitude based on the fundamental Cartesian-coordinate system $x_1, \cdots, x_{3n}$. Let ψ_q denote a probability amplitude which describes the same state of the same assemblage of dynamical systems in terms of a second set of independent variables $q_1, \cdots, q_\lambda$. The transformation of ψ_x into ψ_q can be symbolized by the equation

$$\psi_q = T^{\left(q\atop x\right)}\psi_x,$$

where $T^{\left(q\atop x\right)}$ denotes a linear operator which transforms a function of the coordinates $x_1, \cdots, x_{3n}$ into a function of the eigenvalues of the q coordinates, say $q_1', \cdots, q_\lambda'$.[1] It is, of course, unnecessary that the number of independent variables in the q scheme of coordinates shall be the same as that in the x scheme. The operators employed hitherto for dynamical variables transform functions of the x coordinates into new functions of the same coordinates. An operator of this type can be written as $\alpha^{\left(x\atop x\right)}$, but the superscript can be omitted where, as hitherto, the context precludes any ambiguity.

The transformation of ψ_x into ψ_q must be reversible if the two functions are to contain equally complete descriptions of the assemblage under consideration. We accordingly define the inverse, or reciprocal, transformation (T^{-1}) by the equations

$$\left.\begin{aligned}\psi_q &= T^{\left(q\atop x\right)}\psi_x = T^{\left(q\atop x\right)}(T^{-1})^{\left(x\atop q\right)}\psi_q, \\ \psi_x &= (T^{-1})^{\left(x\atop q\right)}\psi_q = (T^{-1})^{\left(x\atop q\right)}T^{\left(q\atop x\right)}\psi_x.\end{aligned}\right\} \tag{36·2}$$

The use of the term *probability amplitude* for ψ_q is intended to imply (*cf.* pp. 63, 235–236) that by summing $\psi_q\psi_q^*$ over discrete eigen-

[1] *Cf.* HILBERT, VON NEUMANN, and NORDHEIM, *Math. Annalen* **98**, 1 (1927).

values of the q's and integrating over the continuous ones it is possible to evaluate the probability of any range of eigenvalues in q' space. It is convenient to introduce the symbol $\overline{\sum}_{q'}$ to denote this mixed process of summation and integration. If it

is to be extended over a domain G in the λ-dimensional space in which the q''s are laid out as orthogonal coordinates, we use the symbol $\overline{\sum}_{G}{}_{q'}$. The number obtained by applying the operator $\overline{\sum}_{G}{}_{q'}$ to unity will be referred to as the "volume" of G. In case the process $\overline{\sum}_{q'}$ is to be extended over all of q' space, the subscript G is merely omitted.

If the probability of the domain G is to be $\overline{\sum}_{G}{}_{q'}\psi_q\psi_q{}^*$, it is necessary that $\overline{\sum}_{q'}\psi_q\psi_q{}^*$ shall be unity. The obvious parallelism between this expression and the scalar product (ψ_x,ψ_x) suggests the desirability of defining the scalar product of two different probability amplitudes in q' space, say $\psi_q(q')$ and $\varphi_q(q')$ by

Fig. 15.—Illustrating the operation $\overline{\sum}_{q'}$ in two dimensions. In this simple case the circles represent points on the discrete spectra of q_1 and q_2, the vertical (and horizontal) lines are the loci of points on the discrete spectrum of q_1 and the continuous spectrum of q_2 (and *vice versa*), while the crosshatched region belongs to the continuous spectrum of both variables. The diagram shows the discrete and continuous spectra of q_1 overlapping—an unusual case. The process $\overline{\sum}_{q'}$ reduces to a simple summation for the region G_1, to the sum of three single integrals for G_2, to the sum of single integrals and contributions from discrete points in the case of G_3, and to a double integral for G_4.

$$(\psi_q(q'),\ \varphi_q(q'))_{q'} = \overline{\sum}_{q'}\psi_q(q')\varphi_q(q')^*. \tag{36·3}$$

This type of scalar product shares the properties listed in Sec. 22 for the scalar product of two functions of the Cartesian coordinates and reduces to the standard forms (22·16) and (22·13) in special cases.

If the operator $T^{\left(\frac{q}{x}\right)}$ is to convert a properly normalized probability amplitude in x space into a properly normalized probability amplitude in q' space, it is necessary that $(\psi_x,\psi_x)_x = (T^{\left(\frac{q}{x}\right)}\psi_x,\ T^{\left(\frac{q}{x}\right)}\psi_x)_{q'}$. From the linearity of the operator $T^{\left(\frac{q}{x}\right)}$, it follows that if the functions $\psi_x(x)$ and $\varphi_x(x)$ are quadratically integrable in x space their transforms by $T^{\left(\frac{q}{x}\right)}$ have an absolutely convergent scalar product in q' space, such that

$$(T^{\left(\frac{q}{x}\right)}\psi_x(x),\ T^{\left(\frac{q}{x}\right)}\varphi_x(x))_{q'} = (\psi_q(q'),\ \varphi_q(q'))_{q'} = (\psi_x(x),\ \varphi_x(x))_x. \tag{36·4}$$

It is customary to define a *unitary* operator as a linear operator of the type $\alpha = \alpha^{\left(x\right)}_{x}$ which has an adjoint which is also its reciprocal.[1] Let $U^{\dagger}$ denote the adjoint of such a unitary operator U with respect to the linear manifold C and the domain of integration M (*cf.* Sec. 32*d*). If $\psi_1(x)$ and $\psi_2(x)$ belong to C, it follows from the definition of the adjoint that

$$(\psi_1, \psi_2) = (U^{\dagger} U \psi_1, \psi_2) = (U \psi_1, U \psi_2),$$

the domain of integration for the scalar products being M. Thus U, like $T^{\left(q\right)}_{x}$, effects a reversible transformation which *preserves scalar products*. In fact, if U is a reversible linear operator which maps the domain M on itself and preserves scalar products over that domain, it follows at once that it is unitary in the above sense. Hence it is convenient to generalize the definition as follows: *An operator* $T^{\left(q\right)}_{x}$ *is said to be unitary if it has an inverse* $(T^{-1})^{\left(x\right)}_{q}$ *and preserves scalar products in the sense of* (36·4). This definition makes all the operators which transform probability amplitudes from one scheme to another unitary, and is equivalent to our initial definition when applied to operators of the type $\alpha^{\left(x\right)}_{x}$. The transformation produced by a unitary operator is called a *unitary transformation*.

Let $\alpha^{\left(x\right)}_{x}$ and $\alpha^{\left(q\right)}_{q}$ denote the forms of the operator α appropriate to the x and q coordinate systems, respectively. Let $\psi_x(x)$ and $\psi_q(q')$ denote, respectively, an eigenfunction of α in Cartesian-coordinate space and its transform by $T^{\left(q\right)}_{x}$. Consistency demands that

$$\alpha^{\left(x\right)}_{x} \psi_x(x) = \alpha' \psi_x(x); \qquad \alpha^{\left(q\right)}_{q} \psi_q(q') = \alpha' \psi_q(q').$$

It follows that

$$\begin{aligned}
\alpha^{\left(q\right)}_{q} \psi_q(q') &= \alpha' T^{\left(q\right)}_{x} \psi_x(x) = T^{\left(q\right)}_{x} \alpha^{\left(x\right)}_{x} \psi_x(x) \\
&= T^{\left(q\right)}_{x} \alpha^{\left(x\right)}_{x} (T^{-1})^{\left(x\right)}_{q} \psi_q(q').
\end{aligned}$$

In other words, α transforms according to the rule

$$\alpha^{\left(q\right)}_{q} = T^{\left(q\right)}_{x} \alpha^{\left(x\right)}_{x} (T^{-1})^{\left(x\right)}_{q}. \tag{36·5}$$

Such an operator transformation is said to be *canonical*.

The transformation of the Hamiltonian operator H in Cartesian coordinates to the form H_q given in Eq. (35·11) for any system of generalized positional coordinates q_1', q_2', $\cdots$ is a special case of (36·5) in which $T^{\left(q\right)}_{x}$ is the operation of multiplying the function ψ_x to be operated

[1] *Cf., e.g.,* VON NEUMANN, *M.G.Q.,* p. 50.

on by the square root of the Jacobian $\left|\dfrac{\partial(x_1, x_2, \cdots)}{\partial(q_1', q_2', \cdots)}\right|$ of the trans-
formation—the g of Sec. 35 is the *square* of this Jacobian—and sub-
sequently replacing each of the Cartesian coordinates by the equivalent
function of the q''s.

An important property of canonical transformations is that they
carry Hermitian operators in one type of coordinate space over into
Hermitian operators in another. Let us assume, for example, that
$\alpha^{\binom{x}{x}}$ is Hermitian with respect to class D. Let ψ_A and ψ_B denote any two
class D functions in x space. Then

$$(\alpha^{\binom{x}{x}}\psi_A, \psi_B)_x = (\psi_A, \alpha^{\binom{x}{x}}\psi_B)_x.$$

But

$$(\alpha^{\binom{q}{q}}T^{\binom{q}{x}}\psi_A, T^{\binom{q}{x}}\psi_B)_{q'} = (T^{\binom{q}{x}}\alpha^{\binom{x}{x}}(T^{-1})^{\binom{x}{q}}T^{\binom{q}{x}}\psi_A, T^{\binom{q}{x}}\psi_B)_{q'}$$
$$= (T^{\binom{q}{x}}\alpha^{\binom{x}{x}}\psi_A, T^{\binom{q}{x}}\psi_B)_{q'} = (\alpha^{\binom{x}{x}}\psi_A, \psi_B)_x.$$

Similarly,

$$(T^{\binom{q}{x}}\psi_A, \alpha^{\binom{q}{q}}T^{\binom{q}{x}}\psi_B)_{q'} = (\psi_A, \alpha^{\binom{x}{x}}\psi_B)_x.$$

Therefore,

$$(\alpha^{\binom{q}{q}}T^{\binom{q}{x}}\psi_A, T^{\binom{q}{x}}\psi_B)_{q'} = (T^{\binom{q}{x}}\psi_A, \alpha^{\binom{q}{q}}T^{\binom{q}{x}}\psi_B)_{q'}, \qquad (36\cdot6)$$

which means that $\alpha^{\binom{q}{q}}$ is Hermitian with respect to q' space and the linear
manifold of functions obtained by applying $T^{\binom{q}{x}}$ to class D.

The essential thing to be noted here is that, although a single operator
acting on functions of the Cartesian coordinates generates a dynamical
variable, there is a formal difficulty about identifying the variable with
the operator, owing to the fact that the latter has many equivalents
in other probability-amplitude schemes. This difficulty is met by agree-
ing that all operators derived from a single parent by means of (36·4)
shall be regarded as different forms of one and the same operator. To
avoid ambiguity an orthogonal Cartesian-coordinate system will be
assumed throughout the following discussion where there is no definite
statement to the contrary.

36d. Type 1 Operators as Dynamical Variables.—As indicated in
Sec. 36a, it is customary to apply the term *dynamical variable*[1] in quantum

---

[1] The term *observable* was introduced by Dirac in the first edition of his book
(*P.Q.M.*, p. 25) with a meaning analogous to our term *dynamical variable*, but referring
to a particular instant of time. In the second edition the terms observable and
dynamical variable are both used in the same sense as the latter term is used by us.
We also make use of the term observable, but apply it to the restricted class of dynam-

mechanics not only to linear Hermitian[1] operators which are associated with actual or conceptual schemes of observation but also to other operators having similar mathematical properties, but not associated with such methods of measurement.   In effect we include among the dynamical variables all operators which permit the determination of an eigenvalue spectrum and allow us to evaluate a distribution function which would be appropriate to the prediction of the probability of the different eigenvalues for a given state function $\psi(x)$ if a method of measurement were known.   In defining the dynamical variables of quantum mechanics it is convenient to draw an initial distinction between those operators for which the eigenvalue-eigenfunction problem in Cartesian-coordinate space can be given the conventional form associated with (36·1) and those which, like the multiplication operators for the positional coordinates, require a modified formulation of the eigenvalue-eigenfunction problem.   We designate the former as of "type 1" and the latter as of "type 2."   The operator for a given dynamical variable can shift from type 1 to type 2, or *vice versa*, as we change from one scheme of probability amplitudes to another, but so long as we stick to functions of the rectangular coordinates, ψ_x, it is permissible to refer to the dynamical variables themselves as of type 1 or type 2, as the case may be.   We concentrate our attention for the present on the first type.

The degeneracy of the eigenvalues and consequent ambiguity of their eigenfunctions makes it difficult to formulate the properties of the operators singly.   In the cases so far studied, however, the type 1 operators can be united into groups having complete systems of simultaneous eigenfunctions which are nondegenerate.   Each of these eigenfunctions is associated with a corresponding set of eigenvalues and is uniquely determined by that set except for the usual constant of proportionality.   By choosing this factor once and for all in some convenient way consistent with the usual normalization rules, we obtain a perfectly definite system of simultaneous eigenfunctions of the operators of the group and a corresponding uniquely defined expansion theorem.   In the case of a single particle in three dimensions, for example, the operators p_x, p_y, p_z form such a group.   The arbitrary set of eigenvalues p_x', p_y', p_z' has no eigenfunctions except multiples of $e^{\frac{2\pi i}{h}(xp_x' + yp_y' + zp_z')}$.

---

ical variables which are symmetrical functions of the coordinates and momenta of identical particles (*cf.* Sec. 40*d*, p. 310, and Sec. 42*b*, page 339).

Of course it is really illogical to apply the term dynamical variable to something which is not a function of the time for a dynamical system.   It is no more illogical, however, than to apply the term atom, as we do, to systems of particles which are anything but indivisible.   Our excuse is that we have to do with a necessary generalization of the class of quantum analogues of the classical dynamical variables.

[1] The Hermitian property is convenient but not essential.   The less stringent requirement that each operator α shall have an adjoint $\alpha^\dagger$ with which it commutes gives a possible alternative restriction which allows complex eigenvalues.

The operators H, $\mathcal{L}^2$, $\mathcal{L}_z$ form a similar group for the two-particle problem, each set of eigenvalues having an essentially unique eigenfunction $\psi_{nlm}(x,y,z) = R_{nl}(r)\Theta_{lm}(\theta)\Phi_m(\varphi)$.

Not all pairs of operators can be united into such a group, as most pairs are incompatible in the sense that they have no simultaneous eigenfunctions. The mathematical conditions of compatibility will be discussed in Sec. 37. Here it will suffice to note that they are met in the above cases and to postulate that the operators in which we are interested can always be united into such groups. In case an operator α has a purely discrete eigenvalue spectrum, no postulate is necessary as the point is readily proved. Let ψ_{1k}, ψ_{2k}, $\cdots$ denote an orthonormal set of eigenfunctions of α with the common eigenvalue a_k. (If α has a complete system of eigenfunctions, we can form a complete orthonormal set ψ_{nk} in an infinite number of ways.) Let us now define the linear operator β by the equations

$$(\beta\psi,\psi_{nk}) = b_n(\psi,\psi_{nk}) \qquad n, k = 1, 2, 3, 4, \cdots$$

or the equivalent

$$\beta\psi = \sum_k \sum_n b_n(\psi,\psi_{nk})\psi_{nk}.$$

The constants b_n are the eigenvalues of β, which is Hermitian provided they are real. The functions ψ_{nk} are simultaneous eigenfunctions which are nondegenerate since no two of them have the same pair of eigenvalues. This proves that α can be united with at least one other Hermitian operator, *viz.*, β, to form a group with a complete system of nondegenerate simultaneous eigenfunctions.

In order to put the above statements on an exact basis we need unique definitions of the terms eigenfunction and eigenvalue for a general linear operator α. This need confronts us once more with the question of boundary conditions.

In Chap. IV the discrete eigenfunctions of Sturm-Liouville equations with singular boundary points were picked out as solutions of these equations which satisfy the singular-point boundary conditions and hence belong to a linear manifold of functions with respect to which the Sturm-Liouville operator is Hermitian. The discrete eigenfunctions of the Schrödinger equation $H\psi = E\psi$ were identified with class D solutions of that equation, *i.e.*, with solutions which belong to a narrowly defined linear manifold of functions with respect to which H is Hermitian. The continuous-spectrum eigenfunctions in turn were defined as bounded solutions of $H\psi = E\psi$ which are not quadratically integrable but conform to the class D conditions at all finite points. In Sec. 32j we found it necessary to introduce the postulate that the eigendifferentials formed from the continuous-spectrum eigenfunctions belong to a linear manifold, including D, with respect to which H is Hermitian.

If it were possible, one would prefer to use the class D boundary-continuity conditions to define eigenfunctions in all cases, but this procedure does not work. The eigendifferentials are not of this class in cases where the variables are separable, and in the case of some operators ordinarily classed as generators of dynamical variables, and having discrete eigenvalues, no class D eigenfunctions exist. Hence we must use a wider class than D. We accordingly require that the discrete eigenfunctions and the eigendifferentials shall belong to a linear manifold of functions with respect to which the operator in question is Hermitian.

It is convenient at this point to introduce the term *Hermitian manifold of* α for a linear manifold of functions with respect to which the operator α is Hermitian when the domain of integration is properly specified. Unless there is a definite indication to the contrary, we shall hereafter assume that the domain of integration is all Cartesian-coordinate space.

Unfortunately it is possible for an operator to have two or more Hermitian manifolds which cannot be united without destroying the Hermitian property. In such cases the operator can have two or more independent complete systems of eigenfunctions with different spectra. This possibility has been demonstrated by von Neumann[1] in the case of the operator $\dfrac{h}{2\pi i} \dfrac{\partial}{\partial x}$ when the domain of integration is finite. It seems probable, however, that any two Hermitian manifolds of α which contain the manifold of physically admissible functions D are capable of union into a single Hermitian manifold. At any rate we assume the existence of a unique Hermitian manifold of the operator α which contains class D and every other Hermitian manifold which contains D. Let D_α denote the manifold thus specified. We shall at times refer to it as the *type D Hermitian manifold of* α.

Definition: α' is said to belong to the discrete spectrum of the Hermitian operator α if there exist one or more nontrivial[2] solutions of the equation $\alpha\phi = \alpha'\phi$ which belong to the Hermitian manifold D_α.

Definition: α'' is said to belong to the type 1 continuous spectrum of the Hermitian operator α if there exists a continuous family of nontrivial solutions of $\alpha\phi = \alpha'\phi$ for values of α' in the neighborhood of α'' such that $\eta^{-1} \int_{\alpha''}^{\alpha''+\eta} \phi(x|\alpha')d\alpha'$ belongs to the domain D_α and does not vanish identically when η is made arbitrarily small.[3]

---

[1] *Cf. M.G.Q.*, p. 79. In comparing statements in this reference with those made here the reader is cautioned to note the difference between his method of specifying an operator and ours (see footnote 1, p. 263, below).

[2] *I.e.*, solutions which do not vanish identically.

[3] The vertical bar in the symbol $\phi(x|\alpha')$ is used here and will be used hereafter to separate two arguments, or two sets of arguments, of the function in question which belong to different classes.

Definition: $\phi(x_1, \cdots ,x_{3n}|\alpha_1', \cdots , \alpha_\lambda') = \phi(x|\alpha')$ *is a simultaneous eigenfunction of the Hermitian operators* $\alpha_1, \cdots , \alpha_\lambda$ *with the eigenvalues* $\alpha_1', \cdots , \alpha_\lambda'$ *provided that it is a continuous function of all the continuous-spectrum eigenvalues which satisfies each of the equations* $\alpha_k\phi = \alpha_k'\phi$, *and provided that the function* $\Delta_{\alpha'}^\eta\phi$ *derived from* ϕ *by integrating with respect to each of the continuously variable* $\alpha''s$ *over the range of values contained in a small cubical element of side* η *in* α *space containing the point* $\alpha_1', \cdots , \alpha_\lambda'$ *belongs to the type D Hermitian manifold of each of the* $\alpha's$. *The function* $\Delta_{\alpha'}^\eta\phi$ *is then said to be an eigendifferential of the set of* $\alpha's$.

Thus, if α_1', α_2' belong to continuous spectra, while the others belong to discrete spectra, the function

$$\Delta_{\alpha'}^\eta\phi = \int_{\alpha_1'}^{\alpha_1'+\eta}d\alpha_1''\int_{\alpha_2'}^{\alpha_2'+\eta}d\alpha_2''\phi(x|\alpha_1'',\alpha_2'',\alpha_3', \cdots ,\alpha_\lambda') \qquad (36\cdot7)$$

is an eigendifferential provided that it belongs to the type D Hermitian manifold of each α. A discrete eigenfunction may be regarded as a special case of an eigendifferential for which the number of integrations is zero. Formula (36·7) defining $\Delta_{\alpha'}^\eta\phi$ is conveniently generalized to

$$\Delta_{\alpha'}^\eta\phi = \overline{\sum_{G(\alpha',\eta)}}_{\alpha''}\phi(x|\alpha''), \qquad (36\cdot8)$$

where $G(\alpha',\eta)$ is a volume element in α'' space having the form of a "hypercube" with a corner at the point $\alpha_1',\alpha_2', \cdots , \alpha_\lambda'$, and made small enough to include not more than one discrete eigenvalue of any variable. By analogy with (32·25) we adopt the normalization rule

$$(\Delta_{\alpha'}^\eta\phi,\Delta_{\alpha'}^\eta\phi) \equiv \int_\infty \cdots \int|\Delta_{\alpha'}^\eta\phi|^2dx_1 \cdots dx_{3n} = \eta^m, \qquad (36\cdot9)$$

where m denotes the number of continuous-spectrum eigenvalues in the set $\alpha_1', \cdots , \alpha_\lambda'$, *i.e.*, the number of integrations involved in forming $\Delta_{\alpha'}^\eta\phi$. The quantity η^m which forms the right-hand member of (36·9) is seen to be the volume of $G(\alpha',\eta)$ as defined in Sec. 36c. Thus $\eta^{-m}\Delta_{\alpha'}^\eta\phi$ is the average value of $\phi(x|\alpha'')$ for the element $G(\alpha',\eta)$.

The orthogonality of the eigendifferentials of a set of operators which belong to different discrete eigenvalues of some one of them, say α_k, is a consequence of the requirement that these eigendifferentials belong to the Hermitian manifold of α_k (*cf.* p. 206). In view of the work of Sec. 30 we should expect that two eigendifferentials $\Delta_{\alpha'}^\eta\phi$ and $\Delta_{\alpha''}^{\eta'''}\phi$ will also be orthogonal if the ranges for any one of the continuously variable $\alpha's$ do not overlap. A rigorous proof of this proposition based on the Hermitian property is not easily formulated but has been carried through by Carleman.[1] The general orthogonality condition can now be identified with

[1] T. CARLEMAN, *Théorie des Équations Intégrales Singulières à Noyeau Réel et Symmétriques,* Upsala, 1923.

the statement that $(\Delta_{\alpha'}^{\eta'}\phi, \Delta_{\alpha''}^{\eta''}\phi)$ is zero if the regions $G(\alpha',\eta')$ and $G(\alpha'',\eta'')$ used in defining the eigendifferentials in question do not overlap. This includes as special cases all three of the orthogonality conditions (32·26).

If $u(\alpha'')$ is a continuous function of the arguments $\alpha_1'', \alpha_2'', \alpha_3'',$ $\cdots, \alpha_\lambda''$, wherever these arguments are continuously variable,

$$\overline{\sum_{G(\alpha',\eta)}}_{\alpha''} u(\alpha'')\phi(x|\alpha'')$$

is approximately equal to $u(\alpha')\Delta_{\alpha'}^{\eta}\phi$ for small values of η. Hence we may plausibly assume that

$$\left(\overline{\sum_{G}}_{\alpha''} u(\alpha'')\phi(x|\alpha''), \psi\right) = \overline{\sum_{G}}_{\alpha''} u(\alpha'') \lim_{\eta\to 0} [\eta^{-m}(\Delta_{\alpha''}^{\eta}\phi, \psi)]. \quad (36·10)$$

On this assumption it follows that

$$\lim_{\eta'\to 0} \left[(\eta')^{-m}\left(\overline{\sum_{\infty}}_{\alpha''} u(\alpha'')\phi(x|\alpha''), \Delta_{\alpha'}^{\eta'}\phi\right)\right] = u(\alpha'). \quad (36·11)$$

Furthermore,

$$\left(\overline{\sum_{G}}_{\alpha''} u(\alpha'')\phi(x|\alpha''), \overline{\sum_{G}}_{\alpha'} v(\alpha')\phi(x|\alpha')\right) = \overline{\sum_{G}}_{\alpha'} u(\alpha')v^*(\alpha'). \quad (36·12)$$

Equation (36·12) may be regarded as an extended form of the normalization-orthogonality condition to be applied to the eigenfunctions of type 1 operators.

Let the Fourier coefficient $c(\alpha')$ for the arbitrary quadratically integrable function $\psi(x)$ be defined by

$$c(\alpha') = \lim_{\eta\to 0} [\eta^{-m}(\psi(x), \Delta_{\alpha'}^{\eta}\phi)]. \quad (36·13)$$

In case the eigenvalues in the set α' are all discrete, the function $\varphi(x|\alpha')$ is quadratically integrable, and (36·13) reduces to

$$c(\alpha') = (\psi(x), \phi(x|\alpha')). \quad (36·14)$$

A similar reduction can be made if $\psi(x)$ is absolutely integrable—a class D function, for example—while ϕ is bounded. Equations (36·11) and (36·12) now yield the inequality

$$\left(\psi - \overline{\sum_{K}}_{\alpha} c(\alpha')\phi, \psi - \overline{\sum_{K}}_{\alpha'} c(\alpha')\phi\right) = (\psi,\psi) - \overline{\sum_{K}}_{\alpha'} |c(\alpha')|^2 \geqslant 0,$$

where K is any finite region in α' space involving only a finite number of discrete eigenvalues of any α.

The simultaneous eigenfunctions $\phi(x|\alpha')$ are said to form a complete system if

$$(\psi,\psi) = \overline{\sum_{\alpha'}} |c(\alpha')|^2 \quad (36·15)$$

when $\psi(x)$ is any quadratically integrable function of the Cartesian coordinates, the operation $\overline{\sum}_{\alpha'}$ being extended over all α' space. This relation is seen to be equivalent to the statement that we can approximate $\psi(x)$ as closely as desired in the least-squares sense by means of an expression of the form $\overline{\sum}_{\alpha'}\phi(x|\alpha')c(\alpha')$. As a corollary on (36·15) it follows that if ψ_A and ψ_B are two different quadratically integrable functions with the Fourier coefficients c_A and c_B, respectively,

$$(\psi_A, \psi_B) = \overline{\sum}_{\alpha'} c_A(\alpha') c_B(\alpha')^*. \tag{36·16}$$

This result can be derived by the method employed in the proof of (22·32).

The reader will recognize in (36·16) an extrapolation of (36·12). The validity of (36·16) for any pair of eigendifferentials $\Delta^{\nu'}_{\alpha'}\phi$, $\Delta^{\nu''}_{\alpha''}\phi$ is evidently a direct consequence of the normalization-orthogonality relation for eigendifferentials.

We are now in a position to lay down a satisfactory definition of the type 1 operators whose properties were roughly prescribed on p. 249.

Definition: A type 1 operator $\alpha = \alpha^{\binom{x}{x}}$ *is a linear operator which converts functions of the Cartesian coordinates into functions of the Cartesian coordinates and has the following properties:*

(a) α *has a Hermitian manifold* D_α *which includes the class of physically admissible functions* D *and every Hermitian manifold which contains* D.

(b) α *either has a system of nondegenerate eigenfunctions which are complete in the sense of Eqs. (36·15) and (36·16), or can be united with one or more additional operators of the same type to form a set which has a complete system of nondegenerate simultaneous eigenfunctions.* [An eigenfunction is understood to be nondegenerate when there is no other eigenfunction with the same set of eigenvalues which is linearly independent of it. When normalized, it is uniquely determined except for an arbitrary constant factor of absolute value unity (phase factor).]

Any two operators which can enter into a common set of the above sort are said to be mutually compatible. When the simultaneous eigenfunctions of any set of mutually compatible operators are actually nondegenerate, the set is said to be *complete* (*cf.* p. 287). It is to be observed that the completeness of a set of mutually compatible operators is wholly different from the completeness of a system of functions.

The set of dynamical variables to which such a complete set of mutually compatible type 1 operators gives rise constitutes what we shall call a *type 1 coordinate system*.

As previously indicated (Sec. 25b) the root-mean-square convergence of $\overline{\sum}_{\alpha'} c(\alpha')\phi(x|\alpha')$ on the function $\psi(x)$, *i.e.*, the validity of (36·15), is by

no means sufficient for a rigorous proof that the series converges at every point on the function $\psi(x)$. Furthermore the point by point expansion

$$\psi(x) = \overline{\sum}_{\alpha'}\phi(x|\alpha')c(\alpha') \tag{36·17}$$

is not absolutely necessary in quantum mechanics provided that the relation (36·15) is available.[1]   However, this expansion, or an equivalent, is essential to the development of the formalism of the Dirac-Jordan transformation theory.   Hence it is a source of satisfaction to know that the investigations of Weyl in one dimension (Sec. 30b) strongly suggest that the right-hand member of (36·17) does converge uniformly to the value $\psi(x)$ provided that the latter function is continuous and belongs to the Hermitian domain of the α's.   If this suggestion is correct, the expansion (36·17) is good for all physically admissible functions $\psi(x)$ when the α's form a type 1 coordinate system.

A more cumbersome alternative method for calculating $\psi(x)$ from the Fourier coefficients $c(\alpha')$ is to use the formula

$$\psi(x) = \lim_{\eta \to 0}\left[\overline{\sum}_{\alpha'}\phi_\eta(x|\alpha')c(\alpha')\right], \tag{36·18}$$

where $\phi_\eta(x|\alpha')$ is the mean value of $\phi(x|\alpha')$ over a hypercube of side η in x space and containing the point in question.   It is not difficult to show that this equation is a direct consequence of (36·15) if $\psi(x)$ is continuous, even if the series (36·17) is not uniformly convergent.

For this purpose it is convenient to make use of "step functions" defined as follows. Let $w(x)$ be any quadratically integrable function of the Cartesian coordinates x_1, $\cdots$, x_{3n}. Let $w_\eta(x)$ be a function derived from $w(x)$ by dividing coordinate space into hypercubes of side η and replacing $w(x)$ at each point by its average value in the hypercube to which the point belongs.   We call the "smoothed" function $w_\eta(x)$ obtained in this manner the step function of $w(x)$ for the given system of hypercubes because in one dimension the graph of $w_\eta(x)$ would have step form.   I et $\Delta_\eta(x)$ denote the difference $w(x) - w_\eta(x)$.   The scalar product of $\Delta_\eta(x)$ and $w_\eta(x)$ is zero since the contribution of every individual hypercube to that scalar product is zero.   Hence

$$(w(x),\ w(x)) = (w_\eta(x),\ w_\eta(x)) + (\Delta_\eta(x),\ \Delta_\eta(x)).$$

Consequently

$$\|w(x)\| \geqslant \|w_\eta(x)\|. \tag{36·18a}$$

Let us now introduce step functions $\psi_\eta(x)$ and $\phi_\eta(x|\alpha')$ formed from $\psi(x)$ and $\phi(x|\alpha')$ with the same system of hypercubes.   Finally, let $F_\eta(x,R)$ denote the step function formed from the difference

$$F(x,R) = \psi(x) - \overline{\sum}_{\substack{\alpha' \\ G}}\phi(x|\alpha')c(\alpha'),$$

---

[1] It does not appear in the von Neumann formulation of the theory to be described briefly in Sec. 36f.

where G is identified with the interior of a hypersphere of radius R in α' space. Clearly

$$F_\eta(x,R) \; = \; \psi_\eta(x) \; - \; \overline{\sum_{G}}_{\alpha'} \phi_\eta(x|\alpha')c(\alpha').$$

From the completeness relation (36·15) we know that $\lim\limits_{R \to \infty} \|F(x,R)\| = 0$. It follows from the inequality (36·18a) that $\lim\limits_{R \to \infty} \|F_\eta(x,R)\| = 0$. But $F_\eta(x,R)$ is piece-by-piece continuous. Consequently $\lim\limits_{R \to \infty} \|F_\eta(x,R)\| = 0$ implies $\lim\limits_{R \to \infty} F_\eta(x,R) = 0$. Thus

$$\psi_\eta(x) \; = \; \overline{\sum}_{\alpha'} \phi_\eta(x|\alpha')c(\alpha').$$

If the hypercube side η is now allowed to approach zero, the continuity of $\psi(x)$ demands that $\lim\limits_{\eta \to 0} \psi_\eta(x) = \psi(x)$. This proves (36·18).

If $\psi(x)$ is not continuous, we have only to replace the left-hand member of (36·18) by $\lim\limits_{\eta \to 0} \psi_\eta(x)$, which can differ from $\psi(x)$ only at points of discontinuity and must have the same Fourier coefficients as $\psi(x)$. For physical purposes there is no need to distinguish between $\psi(x)$ and $\lim\limits_{\eta \to 0} \psi_\eta(x)$, for every physical prediction made from a wave function involves a process of integration with respect to which they are equivalent.

The above justification of (36·18) suggests a valuable reinterpretation of (36·13). The quantity $\eta^{-m}\Delta_{\alpha'}^m\phi$ which appears in the right-hand member of (36·13) can be identified with a step function formed from $\phi(x|\alpha')$ by introducing hypercubes in α' space and replacing ϕ in each hypercube by the corresponding mean value. The parallelism between (36·13) and (36·18) then becomes complete.

36e. Calculation of Probabilities.—Let us turn our attention next to the physical interpretation of the functions $\phi(x|\alpha')$ and $c(\alpha')$ which appear in (36·17). We shall refer to each of the functions $\phi(x|\alpha')$ and to any discrete-continuous linear combination of these functions involving but a single value of α_k' as an eigenfunction of α_k. Class D eigenfunctions of α_k, if any, are interpreted as descriptions of physically possible subjective states with a unique value of the variable α_k, viz., α_k'. We assume that a measurement of α_k for any member of an assemblage of identical systems in such a state must necessarily yield the single result α_k'. In general the eigenfunctions of α_k, whether discrete or of the continuous-spectrum type, are not of class D. According to our basic assumption they do not represent physically realizable subjective states, but each can be regarded as the limit of a sequence of physically admissible functions. Hence we shall refer to them as the wave functions of states with unique values of α_k, even though we can never realize them exactly in practice.

A simultaneous eigenfunction of two or more variables must then represent a state in which each of these variables has a unique value.

When a complete system of such states exists, there is no theoretical reason, like the Heisenberg inequality of Sec. 16, why simultaneous observations of arbitrary accuracy should not be made on all the variables in the group. Hence it is usual to speak of mutually compatible dynamical variables as *simultaneously measurable*, although we may have no concrete plan for carrying out a simultaneous measurement.

In Sec. 15 we used an operational definition of linear momentum to develop a scheme for calculating the relative values of different measured values of the momentum. The scheme has been generalized for the prediction of energies, and we here carry the generalization a step farther, postulating that the probability that the variables α_1, α_2, $\cdots$, α_λ, which make up a complete set of mutually compatible operators, have a set of values in the domain G of α' space is $\overline{\sum}_{\alpha'} |c(\alpha')|^2$. Mathematically, this means that $c(\alpha')$ plays the role of probability amplitude for the "coordinates" α_1, α_2, $\cdots$, α_λ. In the notation of Sec. 36c $c(\alpha')$ becomes $\psi_\alpha(\alpha')$. Equations (36·13) and (36·14) give explicit form to the operator $T^{(\alpha)}_{(z)}$, while (36·18) does the same thing for $(T^{-1})^{(z)}_{(\alpha)}$.

Physically this postulate means that if a simultaneous exact measurement of all the α's could be carried out for a large number of identical physical systems, so prepared that they have a common initial state described by $\psi(x)$, the most probable value of the fraction yielding measured values of the α's corresponding to points in G is $\overline{\sum}_{\alpha'} |c(\alpha')|^2$.

If we do not know how to carry out such a measurement, we shall still refer to the above expression as the *probability of the domain G in α' space*. In other words, this is our mathematical *definition* of the probability in question. In view of (36·15) it satisfies the fundamental requirement that the sum of the probabilities of all possible sets of simultaneous eigenvalues is unity. Also the probability of a set of discrete eigenvalues is unity in the case of a simultaneous eigenfunction with those eigenvalues.

To get the probability that an individual variable α_k has a value in the range $\alpha_k'' < \alpha_k' < \alpha_k'''$ we have only to identify G with the entire domain in α' space defined by the double inequality, thus summing over all possible values of the other variables in the α coordinate system.

It follows from the above procedure for calculating the probabilities of the eigenvalues of α_k that if the variable α_k has an expectation or mean value for systems in the state $\psi(x)$, that value must be given by the expression

$$\bar{\alpha}_k = \overline{\sum}_{\alpha'} \alpha_k' |c(\alpha')|^2. \tag{36·19}$$

By hypothesis a type 1 operator α_k must be Hermitian with respect to a

linear manifold of functions D_{α_k}, which includes D. It follows that if $\psi(x)$ is a class D function (physically admissible), the expectation value of α_k for the state $\psi(x)$ is also given by the formula

$$\bar{\alpha}_k = (\alpha_k \psi, \psi). \tag{36.20}$$

To prove this proposition we note that if $\psi(x)$ belongs to D_{α_k}, $\alpha_k \psi$ is quadratically integrable so that the scalar product $(\alpha_k \psi, \psi)$ exists and can be evaluated by means of (36.16). Let $d(\alpha')$ denote the general expansion coefficient of $\alpha_k \psi(x)$ with respect to the $\phi(x|\alpha')$'s. With the aid of (36.10) and the Hermitian property we deduce

$$d(\alpha') = \lim_{\eta \to 0} [\eta^{-m}(\alpha_k \psi(x), \Delta_{\alpha'}^\eta \phi(x|\alpha'))]$$

$$= \lim_{\eta \to 0} \left[\eta^{-m}\left(\psi(x), \overline{\sum_{\substack{\alpha'' \\ G(\alpha',\eta)}} \alpha_k'' \phi(x|\alpha'')} \right) \right] = \alpha_k' c(\alpha'). \tag{36.21}$$

By (36.16) the right-hand members of (36.19) and (36.20) are equal, as was to be proved.

The theorem of the equivalence of Eqs. (36.19) and (36.20) is the ultimate justification of the restriction of physically admissible functions to class D and the restriction of dynamical variables to operators Hermitian with respect to class D. A second important inference from. (36.21) is obtained by reverting to the notation of Sec. 36c, in which it takes the form

$$T^{(\alpha)}_{(x)} \alpha_k \overset{(x)}{_{(x)}} \psi(x) = \alpha_k' T^{(\alpha)}_{(x)} \psi(x).$$

It follows at once that

$$\alpha_k \overset{(\alpha)}{_{(\alpha)}} \psi(\alpha') \equiv T^{(\alpha)}_{(x)} \alpha_k \overset{(x)}{_{(x)}} (T^{-1})^{(x)}_{(\alpha)} \psi(\alpha') = \alpha_k' \psi(\alpha'). \tag{36.22}$$

Thus the *operator* $\alpha_k \overset{(\alpha)}{_{(\alpha)}}$ *is the operator which multiplies* $\psi(\alpha')$ *by* α_k'.

In concluding this discussion of type 1 operators we call the reader's attention to the fact that we can define an operator in terms of a complete system of normalized orthogonal eigenfunctions instead of deriving the system of functions from a previously known operator by (36.1). If the functions $\phi(x|\alpha')$ are given, the corresponding operators $T^{(\alpha)}_{(x)}$ and $(T^{-1})^{(x)}_{(\alpha)}$ are defined by Eqs. (36.13) and (36.18). The operator $\alpha_k \overset{(\alpha)}{_{(\alpha)}}$ can be identified with the multiplication operator $[\alpha_k' \times]$ and α_k is then obtained by reversing (36.5). Thus

$$\alpha_k \overset{(x)}{_{(x)}} \psi(x) = (T^{-1})^{(x)}_{(\alpha)} \alpha_k \overset{(\alpha)}{_{(\alpha)}} T^{(\alpha)}_{(x)} \psi(x) = (T^{-1})^{(x)}_{(\alpha)} \alpha_k' c(\alpha').$$

$$= \lim_{\eta \to 0} \overline{\sum_{\alpha'}} \phi_\eta(x|\alpha') \alpha_k' c(\alpha'). \tag{36.23}$$

A plausible generalization of the well-known Fischer-Riesz theorem[1] suggests that for the existence of a quadratically integrable transform $\alpha_k^{\binom{x}{z}}\psi(x)$, it is sufficient as well as necessary that the sum-integral

$$\sum_{/\alpha'}|\alpha_k'c(\alpha')|^2 \tag{36.24}$$

shall exist.

36f. Type 2 Operators as Dynamical Variables; the Method of von Neumann.—

The positional coordinates form the basic dynamical variables of classical theory, and of quantum theory as well. Our original physical interpretation of the quantity $|\Psi|^2 d\tau$ together with the assumption that the wave functions Ψ are defined and continuous over all Cartesian-coordinate space, implies that the spectrum of possible measured values of each Cartesian positional coordinate ranges from $-\infty$ to $+\infty$, and that all of these coordinates are simultaneously measurable to any desired precision. It also defines by implication the range of every classically legitimate positional coordinate $q(x_1, \cdots, x_{3n})$. Finally it tells us that the primary probability amplitude $\psi(x)$ has the same relation to positional measurement that $c(\alpha')$ has to measurements of the α's. Hence we have no immediate need to write down operators for the positional coordinates and work out their eigenfunctions except as a means for unifying the theory of the type 1 dynamical variables with that of the type 2 variables of which the positional coordinates are typical.

If we do wish to unify the theory, we can at once identify the operator for a positional coordinate $q(x_1, \cdots, x_{3n})$ with the operation of multiplying by the number $q(x_1, \cdots, x_{3n})$. We symbolize this operation by $[q\times]$. This identification, suggested on p. 243, is confirmed by Eq. (36.22). Any eigenfunction of the equation

$$q\psi = [q\times]\psi = q'\psi \tag{36.25}$$

must accordingly vanish except at points on the surface

$$q(x_1, \cdots, x_{3n}) = q'. \tag{36.26}$$

Although the operator is linear and Hermitian with respect to the manifold of all quadratically integrable functions with quadratically integrable transforms, it is impossible to set up a corresponding complete normal orthogonal system of eigenfunctions. In fact the solutions of (36.25) are

---

[1] *Cf.* VON NEUMANN, *M.G.Q.*, p. 16; F. RIESZ, *Comptes Rendus*, **144**, 615–619 (1907); E. FISCHER, *Comptes Rendus*, **144**, 1022–1024 (1907). The theorem in the original form given by Riesz for a sequence of real numbers c_n states that if $\sum_{n=1}^{\infty} c_n^2$ is convergent, and if $\phi_1, \phi_2, \cdots$ form a complete orthonormal set of functions, there exists a quadratically integrable function, say f, such that $c_n = (f, \phi_n)$ for every n.

so discontinuous that the ordinary normalization formulas for continuous-spectrum eigenfunctions cannot be fulfilled. Any attempt to represent an ordinary wave function as a continuous linear combination of eigenfunctions of the operator $[q\times]$ must fail, because any integral of the form $\int a(q')\psi_{q'}(x)dq'$ must vanish due to the circumstance that for any fixed set of values of the Cartesian coordinates the integrand must vanish except at the isolated value of q' defined by (36·26).

We shall briefly describe two schemes for dealing with the above difficulty. The first and more rigorous method is due to von Neumann.[1] It involves a reformulation of the eigenvalue-eigenfunction problem which permits us to deal with all dynamical variable operators on the same basis. The von Neumann formulation rests upon the observation that in the case of a real dynamical variable α of either type 1 or type 2, it is possible to resolve any function $\psi(x)$ in the Hermitian manifold of α into the sum of two parts, say ψ_1 and ψ_2, each belonging to the same Hermitian manifold and such that ψ_1 is associated with the portion of the spectrum of eigenvalues below an arbitrary value σ, while ψ_2 is associated with the portion above that point. The resolution has the important property that, independent of ψ,

$$(\alpha\psi_1,\psi_1) \leqslant \sigma(\psi_1,\psi_1), \qquad (\alpha\psi_2,\psi_2) > \sigma(\psi_2,\psi_2).$$

By making a number of such cuts we can resolve ψ into any desired number of parts associated with different non-overlapping intervals of the spectrum. (In the case of a multiplication operator $[q\times]$ the eigenvalue spectrum is identified with the totality of the values which the multiplying factor is allowed to take on.)

Let α denote an arbitrary type 1 operator and let β_1, β_2, $\cdots$, β_λ denote a complete set of type 1 operators of which α is a member ($\alpha = \beta_k$). Let $E_\alpha(\sigma)$ denote the operator defined by

$$E_\alpha(\sigma)\psi(x) = \lim_{\eta\to 0}\left[\overline{\sum}_{\substack{\beta' \\ \alpha'\leqq\sigma}}\phi_\eta(x|\beta')c(\beta')\right]$$

$$= \lim_{\eta\to 0}\left\{\overline{\sum}_{\substack{\beta' \\ \alpha'\leqq\sigma}}\phi_\eta(x|\beta')\lim_{\eta'\to 0}[(\eta')^{-1}(\psi(x),\Delta_{\beta'}^{\eta'}\phi)]\right\}. \quad (36·27)$$

Here the notation is that of Eqs. (36·7) to (36·18). The symbol $\overline{\sum}_{\substack{\beta' \\ \alpha'\leqq\sigma}}$

denotes the application of the operator $\overline{\sum}_{\beta'}$ to all that portion of β' space for which $\alpha' \leqq \sigma$. The operators in the one-parameter family so defined are clearly applicable to every quadratically integrable function $\psi(x)$ and must yield quadratically integrable transforms when applied to such a function. It follows from (36·12) that the Fourier coefficient

[1] VON NEUMANN, *M.G.Q.*, Kap. II, Ziff. 6,7,8.

function $c(\alpha')$ for the primary function $u(x,\sigma) = E_\alpha(\sigma)\psi(x)$ is identically zero in the region $\alpha' > \sigma$ but is equal to the Fourier coefficient function for $\psi(x)$ itself in the region $\alpha' \leqq \sigma$. Equation (36·18) in conjunction with the paragraph following the fine print on p. 256 shows that for physical purposes we may consider that a function is defined by its Fourier coefficients. Hence the application of the operator $E_\alpha(\sigma)$ to $E_\alpha(\sigma)\psi(x)$ leaves the function unchanged.

$$E_\alpha(\sigma)[E_\alpha(\sigma)\psi] = E_\alpha(\sigma)\psi. \tag{36·28}$$

By Eq. (36·16),

$$(E_\alpha(\sigma)\psi_A, \psi_B) = \overrightarrow{\sum_{\substack{\beta' \\ \alpha' \leqq \sigma}}} c_A(\beta')c_B(\beta')^*. \tag{36·29}$$

Hence

$$(E_\alpha(\sigma)\psi_A, \psi_B) = (\psi_A, E_\alpha(\sigma)\psi_B). \tag{36·30}$$

Here $\psi_A(x)$ and $\psi_B(x)$ are any two quadratically integrable functions of the Cartesian coordinates. Thus $E_\alpha(\sigma)$ is Hermitian with respect to the manifold of *all* quadratically integrable functions. Operators with the properties defined by Eqs. (36·28) and (36·30) are called *projection operators*.

The operator $E_\alpha(\sigma)$ also has the properties indicated by the following self-explanatory equations:

$$E_\alpha(\sigma'')E_\alpha(\sigma')\psi = E_\alpha(\sigma')E_\alpha(\sigma'')\psi = E_\alpha(\sigma')\psi; \qquad \sigma' \leqq \sigma'' \tag{36·31}$$

$$\|E_\alpha(\sigma')\psi\| \leqq \|E_\alpha(\sigma'')\psi\|; \qquad \sigma' \leqq \sigma'' \tag{36·32}$$

$$\lim_{\sigma \to -\infty} E_\alpha(\sigma)\psi(x) = 0; \tag{36·33}$$

$$\lim_{\sigma \to +\infty} E_\alpha(\sigma)\psi(x) = \psi(x). \tag{36·34}$$

von Neumann calls a family of projection operators which have the properties (36·31) to (36·34) a *resolution of unity*.

None of the properties listed in defining a resolution of unity makes specific connection with the operator α from which $E_\alpha(\sigma)$ was derived. In order to make such a connection in a generalizable form we note that Eq. (36·23) can be written in the form

$$\alpha\psi(x) = \int_{-\infty}^{+\infty} \sigma d[E_\alpha(\sigma)\psi(x)], \tag{36·35}$$

where the right-hand member is a Stieltjes integral.[1] An alternative

---

[1] The Stieltjes integral is a slight generalization of the ordinary Riemann integral frequently employed by physicists. To define an integral of this type over an interval $a < x < b$ of the x axis, we imagine the interval divided into n equal parts $x_\tau < x < x_{\tau+1}$ and set

$$\int_a^b u(x)dv(x) = \lim_{\eta \to \infty} \sum_{\tau=1}^n u(x_\tau)[v(x_{\tau+1}) - v(x_\tau)].$$

If $v(x)$ is differentiable, the Stieltjes integral reduces to the ordinary Riemannian

equivalent scheme for relating α with the resolution of unity $E_\alpha(\sigma)$ is to use the equation

$$(\alpha\psi,\varphi) = \int_{-\infty}^{+\infty} \sigma \, d(E_\alpha(\sigma)\psi,\varphi), \qquad (36\cdot36)$$

which follows directly from (36·16), (36·21), and (36·29) provided that φ is quadratically integrable and ψ belongs to a Hermitian manifold of α which includes the eigendifferentials $\Delta_{\beta'}^{l}\phi$.

Let N_α denote a Hermitian manifold of the above type. As a corollary on (36·36) we have the equation

$$(\alpha\psi,\alpha\psi) = \int_{-\infty}^{+\infty} \sigma^2 d\|E_\alpha(\sigma)\psi\|^2, \qquad (36\cdot37)$$

valid for every ψ in N_α. The right-hand member of this equation must therefore converge for every such function ψ. It can also happen that the right-hand member, which is defined for every quadratically integrable ψ, is convergent for functions not in N_α. The right-hand member of (36·37) plays the part of the sum of the squares of the absolute values of the Fourier coefficient in the Fischer-Riesz theorem (*cf.* footnote 1, p. 259). Hence, by analogy with that theorem, we may assume that if the right-hand member of (36·37) is convergent, there exists a corresponding quadratically integrable function $w(x)$, whose Fourier coefficients are determined by

$$(w,\varphi) = \int_{-\infty}^{+\infty} \sigma \, d(E_\alpha(\sigma)\psi,\varphi). \qquad (36\cdot37a)$$

Now two functions, whose Fourier coefficients with respect to any complete orthogonal system are the same, can differ only at points of Lebesgue measure zero, *i.e.*, at a set of singular points which are of no importance when one forms a definite integral over any domain. As noted at the end of Sec. 36*d* two functions which differ only in this way are indistinguishable for the purposes of wave mechanics. It follows from (36·36) and (36·37a) that if the function $\alpha\psi$ is defined, it can be identified with w. If it is not defined *a priori*, we can define it as identical with w. If φ belongs to N_α we see from (36·30) that

$$(\alpha\psi,\varphi) = \int_{-\infty}^{+\infty} \sigma d(\psi,E_\alpha(\sigma)\varphi) = (\psi,\alpha\varphi).$$

integral $\int_a^b u(x)\frac{dv}{dx}dx$. In (36·35), however, the function $E_\alpha(\sigma)\psi(x)$ is not continuous in σ or differentiable with respect to σ at the discrete eigenvalues of α. In the case of the corresponding formula [*cf.* Eq. (36.38)]

$$[q \times]\psi(x) = \int_{-\infty}^{+\infty} \xi \, d[E_q(\xi)\psi(x)]$$

there is a discontinuity at $\xi = q$ for every value of q.

Thus we can add the function ψ to the Hermitian manifold N_α without spoiling its Hermitian character, provided that the right-hand member of (36·37) is convergent. Let M_α denote the most inclusive Hermitian manifold obtainable in this way by extending N_α. Then M_α is identical with the manifold of all quadratically integrable functions for which the integral $\int_{-\infty}^{+\infty} \sigma^2 d\|E_\alpha(\sigma)\psi\|^2$ is convergent.

Thus far we have defined $E_\alpha(\sigma)$ by the explicit formula (36·37). von Neumann, however, defines the resolution of unity associated with a Hermitian operator by its properties, *i.e.*, by Eqs. (36·28), (36·30) to (36·34), and (36·36). To be precise, he shows that if the operator α has a Hermitian manifold N_α which is everywhere dense in the Hilbert space of all quadratically integrable functions (*cf.* Sec. 32c, p. 201), and if $E_\alpha{}'(\sigma)$ and $E_\alpha{}''(\sigma)$ are two resolutions of unity for which (36·36) holds, provided that ψ belongs to N_α and φ is quadratically integrable, then $E_\alpha{}'(\sigma) = E_\alpha{}''(\sigma)$. In other words the resolution of unity associated with a given operator operating on a given Hermitian manifold of functions is unique. The working out of such a resolution of unity is von Neumann's equivalent of the eigenvalue-eigenfunction problem. At the end of this section it will be proved that $E_\alpha(\sigma)$ defines a spectrum of eigenvalues and in conjunction with a wave function $\psi(x)$ fixes their probabilities.

Contrary to the usage of this book von Neumann includes in the definition of each Hermitian operator α the specification of a definite Hermitian manifold N_α to which alone he considers it applicable.[1] Hence he can say flatly that a given operator has at most one resolution of unity, whereas our definitions admit the possibility of two or more different resolutions associated with different Hermitian manifolds. In order to remove the ambiguity thus introduced into our theory we shall exclude from consideration for physical purposes all Hermitian manifolds not of type D (*cf.* p. 251).

---

[1] We define the Hamiltonian operator H used in the hydrogen-atom problem by

$$H = -\frac{h^2}{8\pi^2\mu}\nabla^2 - \left[\frac{e^2}{r}\times\right].$$

With this operator von Neumann associates a second, say $\hat{H}$, which is identical with H for all functions $\psi(x,y,z)$ which belong to a manifold M_H with respect to which H is Hermitian. $\hat{H}$ is not defined for solutions of $H\psi = E\psi$ which do not belong to M_H, but can be and is defined for all non-differentiable functions which make

$$\int_{-\infty}^{+\infty} \sigma^2 d[\|E_{\hat{H}}(\sigma)\psi\|^2]$$

convergent. Whenever von Neumann speaks of an Hermitian operator he means one of this type. Every solution of the equation $\hat{H}\psi = E\psi$ is a discrete eigenfunction of H, but if we wish to find these solutions we must first investigate solutions of $H\psi = E\psi$ and then inquire which of them belong to the manifold M_H.

There is no difficulty in formulating a resolution of unity for a multiplicative operator $[q\times]$ as well as for an operator of type 1. It is only necessary to presuppose that $q\psi$ is quadratically integrable when ψ belongs to class D. In that case $[q\times]$ does have a type D Hermitian manifold which includes every other Hermitian manifold.

The family of operators $E_q(\sigma)$ defined by

$$E_q(\sigma)\psi(x) = \psi(x), \qquad q(x) \leqq \sigma \atop E_q(\sigma)\psi(x) = 0, \qquad q(x) > \sigma \right\} \tag{36.38}$$

is readily seen to fulfill all the requirements for a resolution of unity Furthermore Eqs. (36·36) and (36·37) remain valid when $[q\times]$ and $E_q(\sigma)$ are substituted for α and $E_\alpha(\sigma)$, respectively, provided only that ψ belongs to the Hermitian manifold of $[q\times]$. Consequently (36·38) solves the problem of defining a resolution of unity for the operator $[q\times]$.

The question now arises whether, or not, the von Neumann eigenvalue problem is solvable not only for multiplicative operators and for operators known *a priori* to be of type 1, but for all operators Hermitian with respect to a manifold of functions which includes class D. This question has not been considered by competent mathematicians, although a somewhat more general form of the existence problem for Hermitian operators has been treated by von Neumann. It appears from his work that we cannot assume the existence of the desired resolution of unity for every Hermitian operator. Hence it is necessary from this standpoint to require that *if a linear operator* $\alpha\binom{x}{x}$ *with a type D Hermitian manifold is to represent a real dynamical variable in Cartesian-coordinate space, a corresponding resolution of unity must exist. We make this the definition of the operator representing a real dynamical variable in Cartesian-coordinate space and define the dynamical variable itself as the class of operators generated by* $\alpha\binom{x}{x}$ *using transformations of the type* (36·5).

It is now possible to define the probability of the eigenvalues of α in an interval $I\colon \sigma' < \alpha \leqq \sigma''$, very simply in terms of $E_\alpha(\sigma)$. Let $F_\alpha(I)$ denote the operator $E_\alpha(\sigma'') - E_\alpha(\sigma')$. $F_\alpha(I)\psi$ is the part of ψ which belongs to the spectrum interval I. In harmony with the rule of p. 257 we identify $\|F_\alpha(I)\psi\|^2$ with the probability of values of α in the interval I for systems in the state ψ. In other words the interval $d\sigma$ has the probability $d\|E_\alpha(\sigma)\psi\|^2$. This definition is evidently in harmony with Eqs. (36·19) and (36·20). It works as well for type 2 operators as for those of type 1. The definitions of discrete and continuous-spectrum eigenvalues are seen to be implicit in the definition of the probability of a range of eigenvalues. If we plot the probability of the range I as a function of the upper limit σ'', we obtain a monotonically increasing function which in general—*i.e.*, for wave functions containing contributions from all parts of the spectrum—has finite discontinuities at

discrete eigenvalues, a positive slope at points on the continuous spectrum, and zero slope at points which are not eigenvalues. Figure 16 llustrates such a plot.

36g. The Method of Dirac and Jordan.—The second common way of unifying the theories of the type 1 and type 2 operators is to make use of a

$$y = \| F_{\alpha k}(1)\psi \|^2 = \int_{\sigma'}^{\sigma''} d\| E_{\alpha k}(\sigma)\psi \|^2$$
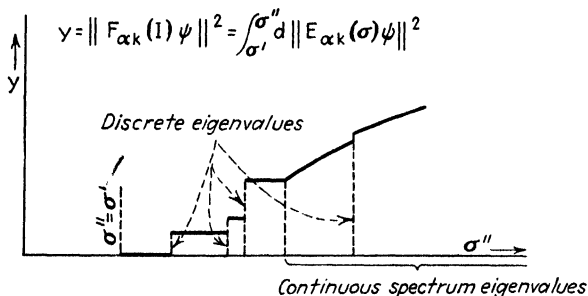
Discrete eigenvalues

Continuous spectrum eigenvalues

FIG. 16.—Plot of $\|F_\alpha(I)\psi\|^2$ against upper limit of range I.

nonrigorous formalism for operators of the second variety which parallels that already developed for operators of the first type.

In developing the procedure we consider first the variable x in one dimension. The function

$$\varphi_{x'}^{(\eta)}(x) = \begin{cases} 0, & \text{if } x \leqq x', \\ \dfrac{1}{\eta}, & \text{if } x' < x \leqq x' + \eta, \\ 0, & \text{if } x' + \eta < x, \end{cases} \tag{36·39}$$

approximates the properties of a discrete eigenfunction of the operator $[x \times]$ if η is a small quantity. It is quadratically integrable, being normalized to the value $\|\varphi_{x'}^{(\eta)}\|^2 = 1/\eta$. It describes a state in which the variable x cannot differ from the mean value $\bar{x} = x' + \frac{1}{2}\eta$ by more than $\frac{1}{2}\eta$. The average value of the square of the error $x - \bar{x}$ is

$$\overline{(x - \bar{x})^2} = \frac{\eta^2}{12}.$$

If we divide the x axis into an infinite number of equal intervals I_τ defined by

$$x_\tau < x \leqq x_{\tau+1}, \qquad x_{\tau+1} - x_\tau = \eta,$$

and set $\varphi_\tau^{(\eta)}(x) = \varphi_{x_\tau}^{(\eta)}(x)$, we obtain an infinite set of mutually orthogonal approximate eigenfunctions. This set is not complete, but if we allow η to approach zero it becomes more and more nearly so, while its members become better approximations to the ideal of a true eigenfunction. Thus, let

$$c^{(\eta)}(x_\tau) = (\psi, \varphi_\tau^{(\eta)}) = \frac{1}{\eta}\int_{x_\tau}^{x_\tau + \eta} \psi(x)dx. \tag{36·40}$$

If we let τ vary with η so that $\lim_{\eta \to 0} x_\tau = x'$, we obtain for the limiting value of the Fourier coefficient $c^{(\eta)}(x_\tau)$:

$$\lim_{\eta \to 0} c^{(\eta)}(x_\tau) = \lim_{\eta \to 0} \left[\int_{-\infty}^{+\infty} \psi(x)\varphi_\tau^{(\eta)}(x)dx \right] = \psi(x'). \qquad (36\cdot41)$$

Then the completeness formula (36·16) has the parallel

$$(\psi_A, \psi_B) = \lim_{\eta \to 0} \left[\sum_\tau c_A^{(\eta)}(x_\tau) c_B^{(\eta)}(x_\tau)^* \eta \right] = \int_{-\infty}^{+\infty} \psi_A(x') \psi_B^*(x')dx'. \qquad (36\cdot42)$$

Furthermore,

$$(x\psi_A, \psi_B) = \lim_{\eta \to 0} \sum_\tau x_\tau c_A^{(\eta)}(x_\tau) c_B^{(\eta)}(x_\tau)^* \eta. \qquad (36\cdot43)$$

Setting $\psi_A = \psi_B$, we obtain a formula for calculating the mean value of x which parallels the scheme used for type 1 operators.
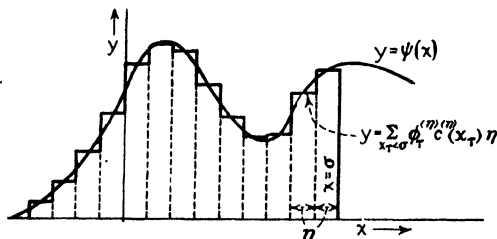

FIG. 17.—Plot of $\psi(x)$ and the step-function approximation to $E_x(\sigma)\psi(x)$.

With the aid of the approximate eigenfunction $\varphi_\tau^{(\eta)}(x)$ we readily derive a form for the projection operator $E_x(\sigma)$ of (36·28) which is similar to that used for type 1 operators. The product $\phi_\tau^{(\eta)}(x)c^{(\eta)}(x_\tau)\eta$ denotes a "rectangular" function of x which vanishes outside the interval

$$x_\tau < x < x_\tau + \eta$$

and inside the interval has an ordinate $c^{(\eta)}(x_\tau)$ equal to the mean ordinate of $\psi(x)$ in the same interval. Taking the point $x = \sigma$ as one of the division points x_τ, we see that $\sum_{x_\tau < \sigma} \phi_\tau^{(\eta)} c^{(\eta)}(x_\tau)\eta$ denotes a step function which approximates $\psi(x)$ for values of x less than σ (cf. Fig. 17).

Let $c^{(\eta)}(x')$ denote the step function whose value in the interval $x_\tau < x' < x_\tau + \eta$ is $c^{(\eta)}(x_\tau)$. Then

$$E_x(\sigma)\psi(x) = \lim_{\eta \to 0} \left[\sum_{x_\tau < \sigma} \phi_\tau^{(\eta)}(x)c^{(\eta)}(x_\tau)\eta \right]$$

$$= \lim_{\eta \to 0} \left[\overline{\sum}_{x' < \sigma}' \phi_{x'}^{(\eta)}(x)c^{(\eta)}(x') \right]. \qquad (36\cdot44)$$

The parallelism with (36·27) is complete. It follows that the probability

of the range $x' < x < x''$ for the wave function $\psi(x)$ is

$$\int_{x'}^{x''} d\|E_x(\sigma)\psi(x)\|^2 = \lim_{\eta \to 0} \left[\sum_{x' < x < x''} |c^{(\eta)}(x)|^2 \right] = \int_{x'}^{x''} |\psi|^2 dx. \quad (36\cdot45)$$

There are sets of continuous approximate eigenfunctions of $[x\times]$ which can be used in similar fashion. Examples are the error function

$$u_n(x - x') = n\pi^{-\frac{1}{2}}e^{-n^2(x-x')^2} \quad (36\cdot46)$$

suggested by Hylleraas[1] and the function

$$w_n(x - x') = \frac{\sin n(x - x')}{(x - x')} \quad (36\cdot47)$$

used in setting up Dirichlet's integral and the Fourier integral theorem. These functions are normalized to the values $n/(2\pi)^{\frac{1}{2}}$ and πn, respectively. They are not orthogonal but become more and more nearly so as n becomes infinite. Thus, if $x' \neq x''$, $\lim_{n \to \infty} (u_n(x - x'), u_n(x - x'')) = 0$. It is not difficult to set up parallels to (36·42), (36·43), and (36·45) with the aid of either of these additional types of approximate eigenfunction.

If any of the sequences of functions $\varphi_{x'}^{(\eta)}(x)$, $u_n(x - x')$, $w_n(x - x')$ were uniformly convergent over the entire interval $-\infty < x < +\infty$ as $1/\eta$, or n, becomes infinite, we should have a corresponding limiting function, say $\delta(x - x')$, with the properties

$$[x\times]\delta(x - x') = x'\delta(x - x'), \quad (36\cdot48)$$

$$\int_{-\infty}^{+\infty} \delta(x - x')dx = \int_{-\infty}^{+\infty} \delta(x - x')dx' = 1, \quad (36\cdot49)$$

$$\int_{-\infty}^{+\infty} f(x)\delta(x - x')dx = f(x'), \quad (36\cdot50)$$

$$\int_{-\infty}^{+\infty} \delta(x - x')\delta(x - x'')dx = \delta(x' - x''). \quad (36\cdot51)$$

Of course the convergence fails at the point $x = x'$ and no function $\delta(x - x')$ with these properties really exists. Nevertheless the existence of a sequence of approximations such as $\varphi_{x'}^{(\eta)}(x)$ is for quantum-mechanical purposes tantamount to the existence of a $\delta(x - x')$. This hypothetical function is called the *Dirac*[2] *δ function* and will be used as it were genuine —a procedure justified by the fact that, if we apply the formulas developed for type 1 operators to Cartesian coordinates and employ eigenfunctions assumed to have the properties (36·48) to (36·51), we get the same results as if the method of von Neumann had been used.

The reader will note that the δ function is not quadratically integrable, as follows from the normalization of the approximation functions, and that it has the type of normalization introduced in Sec. 30 for continuous-spectrum eigenfunctions [*cf.* Eq. (30·20)].

---

[1] E. A. HYLLERAAS, *Grundlagen der Quantenmechanik*, p. 57, Oslo, 1932.
[2] P. A. M. DIRAC, *Proc. Roy. Soc.* **A113**, 621· (1927); *P.Q.M.*, 2d ed., p. 72.

Attention may be called at this point to the fact that the coordinate symbols which form the arguments of the wave functions are not operators but numbers representing possible eigenvalues of the coordinates. The notation will therefore be improved as regards consistency and clarity if we introduce double primes as well as primes for eigenvalues and reserve the unprimed symbols x, y, z for operators. Equation (36·48), for example, is preferably written in the form

$$x\delta(x' - x'') \equiv [x' \times]\delta(x' - x'') = x''\delta(x' - x'').\qquad(36\cdot52)$$

Nevertheless we shall frequently omit the primes on the arguments of the wave functions when no serious ambiguity is likely to result.

*36h. Multiplication Operators in Many Dimensions.—A discussion of the three-dimensional case will perhaps sufficiently illustrate the general problem of multiplication operators in many dimensions. Let the reversible transformation

$$\left.\begin{aligned}
q' &= f(x',y',z'), & x' &= \varphi(q',r',s'),\\
r' &= g(x',y',z'), & y' &= \theta(q',r',s'),\\
s' &= h(x',y',z'), & z' &= \chi(q',r',s'),
\end{aligned}\right\}\qquad(36\cdot53)$$

map x',y',z' space in a one-to-one manner, with the possible exception of singular domains of zero volume, on a domain Q of q',r',s' space. Let the typical multiplication operator q be defined by

$$q = [q' \times] = [f(x',y',z') \times] = f([x' \times],[y' \times],[z' \times]).$$

The eigenfunctions of q in x',y',z' space are then solutions of the equation

$$f(x',y',z')\psi(x',y',z') = q''\psi(x',y',z'),\qquad(36\cdot54)$$

and may be assumed to have the form

$$\delta(q' - q'')F(x',y',z') \equiv \delta[f(x',y',z') - q'']F(x',y',z').$$

Simultaneous eigenfunctions of the three operators q,r,s have the form $\delta(q' - q'')\delta(r' - r'')\delta(s' - s'')F(x',y',z')$, the factor F being chosen to give proper normalization. Here q', r', s' are functions of x', y', z' defined by (36·53).

It will simplify matters if at this point we introduce the simplified Dirac notation for probability amplitudes. Let $(x',y',z'|q'',r'',s'')$ denote the simultaneous eigenfunction in x',y',z' space of the operators q, r, s for the eigenvalues q'', r'', s''. Thus

$$(x',y',z'|q'',r'',s'') = \delta(q' - q'')\delta(r' - r'')\delta(s' - s'')F(x',y',z')\quad(36\cdot55)$$

in adopting this notation let us agree that the density factor (*cf.* p. 239) in the normalization integral shall be unity in x',y',z' space. Conversely we designate a simultaneous eigenfunction of the operators x, y, z in

q'',r'',s'' space by $(q'',r'',s''|x',y',z')$ with the understanding that the density function shall be unity in q'',r'',s'' space. Similarly we write $(x',y',z'|)$ for $\psi(x',y',z')$ and $(q'',r'',s''|)$ for the corresponding probability amplitude in q'',r'',s'' space.

As a normalization condition we use the relation

$$\int dx'dy'dz'(x',y',z'|q'',r'',s'')(x',y,'z'|q''',r''',s''') =$$
$$\delta(q'' - q''')\delta(r'' - r''')\delta(x'' - s''')   (36\cdot56)$$

in analogy with Eq. (36·51). Introducing Eq. (36·55) and transforming the integral over x',y',z' space into an equivalent over q',r',s' space, we have

$$\int_G dq'dr'ds'DF^2\delta(q'-q'')\delta(q'-q''')\delta(r'-r'')\delta(r'-r''')\delta(s'-s'')\delta(s'-s''')$$
$$= \delta(q'' - q''')\delta(r'' - r''')\delta(s'' - s'''),   (36\cdot57)$$

where D is the functional determinant or Jacobian $\partial(x',y',z')/\partial(q',r',s')$. The equation is satisfied if we set $DF^2 = 1$ or identify F with the square root of the Jacobian of the q',r',s' coordinate system with respect to the x',y',z' system. Thus

$$F = \left[\frac{\partial(q',r',s')}{\partial(x',y',z')}\right]^{1/2}.   (36\cdot58)$$

An expansion of the arbitrary wave function $(x',y',z'|)$ into simultaneous eigenfunctions of q, r, and s would have the form

$$(x',y',z'|) = \int_Q (x',y',z'|q'',r'',s'')(q'',r'',s''|)dq''dr''ds''.   (36\cdot59)$$

Such an expansion follows directly from the properties of the δ function if we introduce the symbol Γ for the substitution

$$x' \to \varphi(q',r',s'),   y' \to \theta(q',r',s'),   z' \to \chi(q',r',s'),$$

and define $(q'',r'',s''|)$ by the equations

$$(q'',r'',s''|) = \left[\frac{\partial(x'',y'',z'')}{\partial(q'',r'',s'')}\right]^{1/2}\Gamma(x'',y'',z''|)$$
$$= \int_\infty (q'',r'',s''|x',y',z')(x',y',z'|)dx'dy'dz',   (36\cdot60)$$

$$(q'',r'',s''|x',y',z') = \left[\frac{\partial(x'',y'',z'')}{\partial(q'',r'',s'')}\right]^{1/2}\delta(x' - x'')\delta(y' - y'')\delta(z' - z'').$$
$$(36\cdot61)$$

Equation (36·61) evidently defines a properly normalized simultaneous eigenfunction of the variables x, y, z in q'',r'',s'' space exactly analogous with Eqs. (36·55) and (36·58), while Eqs. (36·59) and (36·60) form parallels to Eq. (36·17) and the Fourier-coefficient formulas (36·13) and (36·14).

The two real functions $(x',y',z'|q'',r'',s'')$, $q('',r'',s''|x',y',z')$ are equivalent and may be set equal to each other, for they yield the same result when multiplied by an arbitrary function of either set of variables and integrated over the corresponding coordinate space.

By direct transformation we readily verify the scalar-product relation [*cf.* Eq. (36·16)]

$$\int_\infty (x',y',z'|A)(x',y',z'|B)\,^*dx'dy'dz'$$
$$= \int_Q (q'',r'',s''|A)(q'',r'',s''|B)\,^*dq''dr''ds'', \quad (36\cdot62)$$

as well as the mean-value theorem

$$\bar{q} = \int_\infty q'(x',y',z')\,|(x',y',z'|)|^2dx'dy'dz' = \int_Q q''|(q'',r'',s''|)|^2dq''dr''ds''.$$
$$(36\cdot63)$$

Thus the functions $(x',y',z'|)$ and $(q'',r'',s''|)$ can be identified, respectively, with the functions ψ_x and ψ_q of Eq. (35·9)

36i. Transformation of Probability Amplitudes from One Arbitrary Coordinate Scheme to Another.—It is now evident that the formalism developed for type 1 coordinate systems can be adapted to include ordinary positional coordinate systems. This can be done either with a system of approximate eigenfunctions, or, less rigorously, with the aid of the δ function. The essential feature of this formalism is that it gives a unitary operator $T^{(\alpha)}_{\left(\begin{smallmatrix}x\end{smallmatrix}\right)}$ which can be used to transform any probability amplitude $\psi_x(x')$ in Cartesian-coordinate space into an equivalent probability amplitude in a space in which the eigenvalues of the new coordinates are laid out "at right angles to one another."

In Eqs. (36·13), (36·18), (36·41), and (36·44) the operator $T^{(\alpha)}_{\left(\begin{smallmatrix}x\end{smallmatrix}\right)}$ and its inverse are expressed as the limits of sum-integrals when η approaches zero. Equations (36·14), (36·17), (36·50), (36·59), and (36·60) give these operators the more convenient form of direct sum-integrals but are of less generality and rigor.

We are thus led to *conjecture that every linear operator* $\alpha^{\left(\begin{smallmatrix}x\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}x\end{smallmatrix}\right)}$ *which conforms to the condition laid down on p. 264—and hence forms the representative in x' space of a real dynamical variable α—can be brought within the scope of the type 1 formalism.* This conjecture implies **(a)** *that the system of eigenvalues of any such operator can be made to form one element of a complete coordinate system* α_1', α_2', $\cdots$, α_λ' *connected with the Cartesian-coordinate system by means of a unitary operator* $T^{(\alpha)}_{\left(\begin{smallmatrix}x\end{smallmatrix}\right)}$; **(b)** *that the operator* $T^{(\alpha)}_{\left(\begin{smallmatrix}x\end{smallmatrix}\right)}$ *and its inverse can be expressed as the limits of corresponding*

sum-integrals;[1] **(c)** *that by means of the δ function or similar devices it is always possible to express the operators* $T^{(\alpha)}_{(x)}$, $(T^{-1})^{(x)}_{(\alpha)}$ *as direct sum-integrals.* These assumptions form an essential basis for any attempt to pass from the von Neumann formulation of quantum-mechanical theory to the Dirac-Jordan transformation theory. So far as the writer is aware no one has carried through the details of a complete proof of the validity of the assumptions **(a)** and **(b)**, but it is possible to go a long way toward formulating such a proof and there is apparently no very good reason to question these items seriously. The third assumption is more doubtful and should be clarified by a more complete specification of the allowed procedure. Nevertheless we shall adopt it as extremely useful for heuristic purposes, while cautioning the reader to check by more rigorous methods any suspicious results obtained from postulate **(c)**.

We proceed to summarize the formulas implied by assumptions **(a)** and **(b)**, recasting them at the same time in the interests of symmetry. Let $\alpha_1, \alpha_2, \cdots, \alpha_\lambda$ denote any set of mutually compatible dynamical variables with a sequence of sets of definite approximate simultaneous eigenfunctions, say $\phi(x'|\alpha')_\eta$ for which the properties of orthogonality and completeness are valid in the limit as the parameter η approaches zero. The probability amplitude in α' space corresponding to the wave function $(x'|) \equiv (x_1', \cdots, x_{3n}'|)$ is

$$(\alpha'|) \equiv T^{(\alpha)}_{(x)}(x'|) = \lim_{\eta \to 0} \int_\infty \phi(x'|\alpha')_\eta^*(x'|)dx_1'dx_2' \cdots dx_{3n}'$$

$$= \lim_{\eta \to 0} \overline{\sum_{x'}} \phi(x'|\alpha')_\eta^*(x'|). \tag{36.64}$$

The reverse transformation is given in all cases of interest, where $(x'|)$ is at least piece-by-piece continuous, by

$$(x'|) = (T^{-1})^{(x)}_{(\alpha)}(\alpha'|) = \lim_{\eta \to 0} \overline{\sum_{\alpha'}} \psi(x'|\alpha')_\eta(\alpha'|). \tag{36.65}$$

Here $\psi(x'|\alpha')_\eta$ is in all cases closely related to $\phi(x'|\alpha')_\eta$, and in Eqs. (36.41) and (36.44)—which becomes a special case of (36.65) when $\sigma = +\infty$ — the two functions are sensibly the same. In case the α's are of type 1 [*cf.* Eqs. (36.13) and (36.18)], the functions $\phi(x'|\alpha')_\eta$ and $\psi(x'|\alpha')_\eta$ are mean values of the exact simultaneous eigenfunction $(x'|\alpha')$ over hypercubes of side η in α' space and x' space, respectively. Since η is ultimately to be set equal to zero, we can in each case substitute for $\phi(x'|\alpha')_\eta$ or $\psi(x'|\alpha')_\eta$, as the case may be, a step function $(x'|\alpha')_\eta$ defined as follows. Let the *combined* space of the α''s and the x''s be divided up into hyper-

---

[1] In the future we shall refer to a unitary operator $T^{(\alpha)}_{(x)}$ which conforms to the assumption **(b)** as a *unitary integral operator*.

cubes of side η. In each hypercube we give $(x'|\alpha')_\eta$ a constant value equal to the average value of $(x'|\alpha')$ in that hypercube, *viz.*,

$$(x'|\alpha')_\eta = \frac{\overline{\sum}_{x',\alpha'}^{\text{Hypercube}}(x'|\alpha')}{\overline{\sum}_{x',\alpha}^{\text{Hypercube}} 1}. \tag{36·66}$$

A similar process is readily carried out for variables of type 2 and we conclude that in general the transformations $T^{\binom{\alpha}{x}}$ and $(T^{-1})^{\binom{x}{\alpha}}$ can be given the symmetrical form[1]

$$(\alpha'|) = T^{\binom{\alpha}{x}}(x'|) = \lim_{\eta\to 0} \overline{\sum}_x (\alpha'|x')_\eta (x'|), \tag{36·64a}$$

$$(x'|) = (T^{-1})^{\binom{x}{\alpha}}(\alpha'|) = \lim_{\eta\to 0} \overline{\sum}_{\alpha'}(x'|\alpha')_\eta (\alpha'|), \tag{36·65a}$$

where $(\alpha'|x')_\eta$ is defined by

$$(\alpha'|x')_\eta = (x'|\alpha')_\eta^*. \tag{36·67}$$

As the Cartesian coordinates have continuous spectra ranging from $-\infty$ to $+\infty$, the operator $\overline{\sum}_{x'}$ in (36·64a) is just a multiple integral. If the operators α also have purely continuous spectra, $\overline{\sum}_{\alpha'}$ in (36·65a) also reduces to a simple multiple integral, but if some or all the α's have discrete spectra it takes the form of a pure summation, or a mixed summation-integration process as the case may be.

Adopting the assumption (c), *i.e.*, interchanging the order of the operations $\lim_{\eta\to 0}$ and $\overline{\sum}_{x'}$, or $\overline{\sum}_{\alpha'}$, as the case may be, we throw the

---

[1] In view of the symmetry of the formulas for $T^{\binom{\alpha}{x}}$ and $(T^{-1})^{\binom{x}{\alpha}}$ it would appear unnecessary to preserve the awkward notation $(T^{-1})^{\binom{x}{\alpha}}$ for the latter, the symbol $T^{\binom{x}{\alpha}}$ being adequate. Such a contraction of notation is not desirable, however, as $T^{\binom{\alpha}{x}}$ is not uniquely defined by the two sets of coordinates involved. If we multiply the functions $(x'|\alpha')_\eta$ by an arbitrary phase factor $e^{iu(\alpha',\eta)}$ which in the limit $\eta = 0$ becomes a continuous function of the α' eigenvalues where they form a continuous set, we obtain a new approximate probability amplitude as good as the original one but yielding a different transformation operator, say $Q^{\binom{\alpha}{x}}$. As indicated in Sec. 36e the physical results derived from one choice of phases is the same as from another, but it is necessary to be consistent and the notation $(T^{-1})^{\binom{x}{\alpha}}$ is a reminder that the choice of phases in Eqs. (36·64a) and (36·65a) must be the same; otherwise one operator would not be the inverse of the other.

transformation equations into the form

$$(\alpha'|) = T^{\binom{\alpha}{x}}(x'|) = \overline{\sum}_{x'}(\alpha'|x')(x'|), \tag{36.68}$$

$$(x'|) = (T^{-1})^{\binom{x}{\alpha}}(\alpha'|) = \overline{\sum}_{\alpha'}(x'|\alpha')(\alpha'|), \tag{36.69}$$

with the understanding that $(\alpha'|x')$ will have the properties of a δ function if one or more of the α's are dynamical variables of type 2. In view of (36.67), it is necessary that

$$(x'|\alpha') = (\alpha'|x')^*. \tag{36.70}$$

Let us now consider a third set of independent variables $\beta_1, \cdots, \beta_\sigma$, with a unitary transformation operator $T^{\binom{\beta}{x}}$. Let $T^{\binom{\beta}{\alpha}}$ denote the operator

$$T^{\binom{\beta}{\alpha}} = T^{\binom{\beta}{x}}(T^{-1})^{\binom{x}{\alpha}}. \tag{36.71}$$

It is linear and has the inverse

$$(T^{-1})^{\binom{\alpha}{\beta}} = T^{\binom{\alpha}{x}}(T^{-1})^{\binom{x}{\beta}}. \tag{36.72}$$

Since

$$(T^{\binom{\beta}{\alpha}}(\alpha'|A), T^{\binom{\beta}{\alpha}}(\alpha'|B))_{\beta'} = ((T^{-1})^{\binom{x}{\alpha}}(\alpha'|A), (T^{-1})^{\binom{x}{\alpha}}(\alpha'|B))_{x'} = ((\alpha'|A), (\alpha'|B))_{\alpha'}, \tag{36.73}$$

$T^{\binom{\beta}{\alpha}}$ is unitary. Finally it transforms the operator $\alpha_k^{\binom{\alpha}{}} = [\alpha_k' \times]$ into $\alpha_k^{\binom{\beta}{}}$ by the rule of Eq. (36.22). Thus $T^{\binom{\beta}{\alpha}}$ has the same general properties as $T^{\binom{\beta}{x}}$ and we should expect to be able to give it a similar form. In fact

$$T^{\binom{\beta}{\alpha}}(\alpha'|) = \lim_{\eta \to 0} \overline{\sum}_{x'}(\beta'|x')_\eta \left[\lim_{\zeta \to 0} \overline{\sum}_{\alpha'}(x'|\alpha')_\zeta(\alpha'|) \right] \tag{36.74}$$

$$= \lim_{\eta \to 0} \overline{\sum}_{\alpha'}(\beta'|\alpha')_\eta(\alpha'|),$$

where

$$(\beta'|\alpha')_\eta = \overline{\sum}_{x'}(\beta'|x')_\eta(x'|\alpha')_\eta = (\alpha'|\beta')_\eta^*. \tag{36.75}$$

Using the postulate (c) we get the simplified equations

$$(\beta'|) = \overline{\sum}_{\alpha'}(\beta'|\alpha')(\alpha'|), \qquad (\alpha'|) = \overline{\sum}_{\beta'}(\alpha'|\beta')(\beta'|), \tag{36.76}$$

which we shall ordinarily assume to hold for any two complete sets of coordinates.

Whether we give the transformation $T^{\binom{\alpha}{x}}$ the form (36.64a) or (36.68) the application of the term *probability amplitude* to $(\alpha'|)$ implies that the probability of the eigenvalues in the region G of α' space for an assemblage of systems in a state described by $(x'|)$ is to be defined as

$$Q(G) = \overline{\sum}_{\alpha'}|(\alpha'|)|^2. \tag{36.77}$$

The probability of an individual discrete eigenvalue α_k', or of an individual elementary range $d\alpha_k'$, are similarly defined as $Q(\alpha_k')$ and $Q(\alpha_k')d\alpha_k'$, respectively, where

$$Q(\alpha_k') = \overline{\sum}_{\alpha' \neq \alpha_k'} |(\alpha_1', \cdots, \alpha_k', \cdots |)|^2. \tag{36.78}$$

Here the symbol $\overline{\sum}_{\alpha' \neq \alpha_k'}$ means the application of the operation $\overline{\sum}$ to the entire range of values of each of the α' coordinates except α_k'.

So far we have made no general statement about the range of applicability of the unitary operators $T^{\binom{\alpha}{x}}$, $(T^{-1})^{\binom{x}{\alpha}}$. If the α's are all of type 1, it follows from the completeness of the system of simultaneous eigenfunctions [*cf.* Eq. (36.15)] that if $\psi(x')$ is quadratically integrable, the transform $T^{\binom{\alpha}{x}}\psi(x')$ exists and is quadratically integrable. Conversely, by the previously assumed extension of the Fischer-Riesz theorem (*cf.* footnote 1, p. 259), and remark following Eq. (36.37) on p. 262), it follows that if the probability amplitude $c(\alpha')$ is quadratically summable its transform $\psi(x') = (T^{-1})^{\binom{x}{\alpha}}c(\alpha')$ exists and is quadratically integrable. In case the α's are multiplication operators, Eq. (36.60) shows that $T^{\binom{\alpha}{x}}$ reduces to the operation of multiplying by the square root of the appropriate Jacobian determinant and substituting for each Cartesian coordinate the equivalent function of the α's. Here again the operators $T^{\binom{\alpha}{x}}$ and $(T^{-1})^{\binom{\alpha}{x}}$ are defined and yield quadratically summable transforms provided that the function operated on is quadratically summable. Hence we shall *assume* that this property holds in the most general case.

Consider next the Hermitian manifold of the real dynamical-variable operator $\alpha = \alpha^{\binom{x}{x}}$. If ψ_A belongs to this manifold, it is necessary that ψ_A and $\alpha\psi_A$ shall be quadratically integrable. Then the transforms of these functions in α' space are $(\alpha'|A)$ and $\alpha'(\alpha'|A)$. If ψ_A and ψ_B satisfy this necessary condition, it follows from Sec. 36c that

$$(\alpha\psi_A, \psi_B) = \overline{\sum}_{\alpha'} \alpha'(\alpha'|A)(\alpha'|B)^* = (\psi_A, \alpha\psi_B).$$

Thus, the necessary condition is sufficient and it follows that the operator α is Hermitian with respect to the manifold of all quadratically integrable functions with quadratically integrable transforms.

We can now prove our right to apply the operator α term by term, or under the integral sign, to the expansions (36.65a) and (36.69), provided ·

that the function expanded, *i.e.* $(x'|)$, has a quadratically integrable transform by α. In that case

$$\alpha(x'|) = (T^{-1})^{\binom{x}{a}}\alpha'T^{\binom{\alpha}{x}}(x'|) = (T^{-1})^{\binom{x}{a}}\alpha'(\alpha'|)$$

$$= \lim_{\eta \to 0}\overline{\sum}_{\alpha'}(x'|\alpha')_\eta\alpha'(\alpha'|) \sim \overline{\sum}_{\alpha'}(x'|\alpha')\alpha'(\alpha'|), \quad (36\cdot79)$$

which establishes the proposition.

In Sec. 32*d* we gave a proof of the reality of the discrete eigenvalues of the standard Hamiltonian operator based on its Hermitian character and immediately applicable to any other Hermitian operator. We have not yet given a general proof of the reality of the continuous-spectrum eigenvalues, though the discussion of the continuous spectrum as the limit of a discrete spectrum in Secs. 30 and 32*j* would lead us to infer that reality. The desired proof is readily established, however, by means of the transformation to an α' space. Thus if $(x'|)$ belongs to the Hermitian manifold of α and $(\alpha'|)$ is its transform by $T^{\binom{\alpha}{x}}$, we have

$$(\alpha(x'|),(x'|)) - ((x'|),\alpha(x'|)) = \overline{\sum}_{\alpha'}[\alpha' - (\alpha')^*]|(\alpha'|)|^2 = 0.$$

This equation must hold for every $(\alpha'|)$ which together with $\alpha'(\alpha'|)$ is quadratically summable. Hence $\alpha' - (\alpha')^*$ must vanish for all values of α', *i.e.*, α' is real.

36j. Dynamical Variables with Complex Eigenvalues.[1]—The definitions given on p. 251 of the Hermitian manifold of an operator α and of its eigenvalues are converted into definitions of the *unitary manifold* of the unitary operator $U = U^{\binom{x}{x}}$ and of its eigenvalues by substituting U for α and *unitary* for *Hermitian* throughout. The discrete eigenfunctions of U are defined as solutions of $U\psi = U'\psi$ which belong to the unitary manifold of U. The orthogonality of the eigenfunctions of such an operator can be proved in the same way as for an Hermitian operator. In fact the whole theory of the transformations $T^{\binom{\alpha}{x}}$, $(T^{-1})^{\binom{x}{a}}$ goes through equally well if we allow the operators α, which make up the coordinate system, to be either unitary or Hermitian. The eigenvalues of a unitary operator, however, are in general complex and of absolute value unity. To prove this we note that if U^{-1} is the inverse of U and ϕ is an eigenfunction of U with the eigenvalue u,

$$U^{-1}U\phi = uU^{-1}\phi = \phi.$$

---

[1] The complex eigenvalues of non-Hermitian operators here introduced must not be confused with the complex eigenvalues of Hermitian operators used in the theory of radioactive disintegration (Sec. 31*f*).

It follows that ϕ is also an eigenfunction of U^{-1} with the eigenvalue u^{-1}. Then

$$(U\phi,\phi) - (\phi,U^{-1}\phi) = [u - (u^{-1})^*](\phi,\phi) = 0.$$

Therefore $(u^{-1})^* = u$, or $|u| = 1$. A more general proof applicable to the continuous spectrum can be worked out with the aid of a transformation to a space in which the eigenvalues of U are coordinates.

It now becomes evident that the real dynamical-variable operators are a part of a wider class of linear operators which can also be thrown into multiplicative form by a suitable canonical transformation but whose eigenvalues are in general complex. Let $T^{\binom{q}{x}}$ denote any unitary integral operator defined for every quadratically integrable function $(x'|)$, and let $f(q')$ denote any real or complex function of the coordinates in the q space defined by $T^{\binom{q}{x}}$. Let γ be the operator in Cartesian-coordinate space defined by

$$\gamma \equiv \gamma^{\binom{x}{x}} = (T^{-1})^{\binom{x}{q}} f(q') T^{\binom{q}{x}}. \tag{36.80}$$

$\gamma(x'|)$ is defined and quadratically integrable if

$$\overline{\sum}_{q'} |f(q') T^{\binom{q}{x}}(x'|)|^2 \tag{36.81}$$

converges. Let $(x'|A)$ and $(x'|B)$ denote two quadratically integrable functions with quadratically integrable transforms $(q'|A)$, $(q'|B)$, respectively. Let $\gamma^\dagger$ be defined by

$$\gamma^\dagger = (T^{-1})^{\binom{x}{q}} f(q')^* T^{\binom{q}{x}}. \tag{36.82}$$

Then

$$(\gamma(x'|A), (x'|B)) = \overline{\sum}_{q'} f(q')(q'|A)(q'|B)^* = ((x'|A), \gamma^\dagger(x'|B)). \tag{36.83}$$

Thus $\gamma^\dagger$ is adjoint to γ with respect to Cartesian-coordinate space and the linear manifold of all functions which make the expression (36·81) convergent. We call this the *adjoint manifold* of γ. Finally, we generalize the term *dynamical variable* to include the class of operators composed of γ and its canonical transforms, provided that the adjoint manifold contains the manifold of physically admissible wave functions D.

The eigenvalues of the general dynamical variable γ are the possible values of $f(q')$ and the eigenfunctions are defined by the transformation operator $T^{\binom{q}{x}}$ when that is thrown into the form (36·64) or (36·68). The eigenfunctions of $\gamma^\dagger$ are identical with those of γ, but the eigenvalues of γ and $\gamma^\dagger$ for any given eigenfunction are complex conjugates. The real dynamical variables form that subclass of the general dynamical

variables which are Hermitian and have real eigenvalues. By substituting a real function for $f(q')$ we can always derive a real dynamical variable with the same system of eigenfunctions from any given complex variable γ. Hence the eigenfunctions of the general dynamical variables must always have the same orthogonality properties as the eigenfunctions of real dynamical variables.

So far as concerns the theory of measurements it would suffice to restrict the discussion to variables with real eigenvalues, but in the mathematical discussion of the symmetry properties of the Schrödinger equation the concept of the broader class of variables is useful.

CHAPTER VIII

COMMUTATION RULES AND RELATED MATTERS

37. SIMULTANEOUS EIGENFUNCTIONS AND THE COMMUTATION OF DYNAMICAL VARIABLES

37a. Operator Algebra.—The sum and product of two linear operators are defined by the equations

$$(\alpha + \beta)\psi = \alpha\psi + \beta\psi, \tag{37.1}$$
$$(\alpha\beta)\psi = \alpha(\beta\psi). \tag{37.2}$$

The multiplication of an operator α by a number c follows the rule

$$(c\alpha)\psi = c(\alpha\psi) = \alpha(c\psi). \tag{37.3}$$

These operations conform to the rules of ordinary algebra except for the commutative law of multiplication and so define an algebra for linear operators similar to the algebra of matrices.

If the functions ψ and χ belong to the adjoint manifold of α,

$$(c\alpha\psi, \chi) = c(\psi, \alpha^\dagger\chi) = (\psi, c^*\alpha^\dagger\chi).$$

Hence $c\alpha$ has the adjoint $c^*\alpha^\dagger$, the adjoint manifold being identical with that of α. If α is a dynamical variable, $c\alpha$ is also a dynamical variable. If α is a real dynamical variable, *i.e.*, is Hermitian, $c\alpha$ is a real dynamical variable if, and only if, c is real.

Let us now assume that α and β are dynamical variables and that ψ and χ belong to class D. It follows that

$$((\alpha + \beta)\psi, \chi) = (\alpha\psi, \chi) + (\beta\psi, \chi)$$
$$= (\psi, \alpha^\dagger\chi) + (\psi, \beta^\dagger\chi) = (\psi, (\alpha^\dagger + \beta^\dagger)\chi).$$

Hence $\alpha + \beta$ has an adjoint manifold which includes class D. To the additional question whether $\alpha + \beta$ gives rise to a solution of the von Neumann eigenvalue problem, or has a complete orthonormal set of eigenfunctions, because α and β separately have that property, we can offer no positive answer. However, von Neumann has shown that Hermitian operators which do not yield such solutions are to be regarded as exceptional, whereas unitary operators always yield corresponding resolutions of unity.[1] Hence it is to be expected that in most cases the appropriate resolution of unity will exist. If this hypothesis is correct, $\alpha + \beta$ is a dynamical variable. Furthermore $\alpha + \beta$ is Hermitian and a real dynamical variable if α and β are Hermitian.

[1] VON NEUMANN, *M.G.Q.*, II, 9.

The case of the product of two dynamical variables involves further uncertainty. We cannot be sure that $\alpha\beta\psi$ will be quadratically integrable if $\alpha\psi$ and $\beta\psi$ are quadratically integrable. However, the severe restrictions on class D require that $\alpha\beta\psi$ and $\beta\alpha\psi$ shall be quadratically integrable provided that ψ is of class D and provided that α and β are positive powers of the Cartesian coordinates or of their conjugate momenta. Let us assume that α and β are dynamical variables of this basic classical type. Then if ψ and χ belong to class D

$$(\alpha\beta\psi, \chi) = (\beta\psi, \alpha\chi) = (\psi, \beta\alpha\chi).$$

Thus $\alpha\beta$ has an adjoint manifold which includes class D. If we make the optimistic assumption that $\alpha\beta$ yields a solution of the von Neumann eigenvalue problem, we may infer that $\alpha\beta$ is a dynamical variable. It is Hermitian if, and only if, the operator $\alpha\beta$ and its adjoint $\beta^\dagger\alpha^\dagger$ are equivalent. If α and β are Hermitian, it suffices that they commute. If they are Hermitian but do not commute, we can always form a Hermitian operator from the products $\alpha\beta$ and $\beta\alpha$ by taking their sum, for

$$((\alpha\beta + \beta\alpha)\psi, \chi) = (\psi, (\alpha\beta + \beta\alpha)\chi).$$

In the limiting case of a sharply defined wave packet the average value of $\frac{1}{2}(\alpha\beta + \beta\alpha)$ can be identified with the classical value of the product of the variables α and β (*cf.* Sec. 13), so that this symmetrized Hermitian operator forms the quantum-mechanical analogue of the classical product variable.

There is evidently a large class of classical dynamical variables which can be correlated with quantum-mechanical dynamical-variable operators. We have no reason to suppose, however, that every function of the coordinates and momenta defining a classical dynamical variable can be correlated with a quantum-mechanical dynamical variable, nor is it altogether certain that when such a correlation exists it is unique.

Theorem: Any algebraic relationship between two or more operators based on Eqs. (37·1), (37·2), and (37·3) is preserved when the operators are subjected to a canonical transformation of the type (36·5). The proof is left to the reader.

37b. Functions of a Single Operator.—Equations (37·1), (37·2), and (37·3) give meaning to the various powers of an operator α and hence to any function of α expressible as a polynomial or power series. Consider the simple case

$$f(\alpha) \equiv c_1\alpha + c_2\alpha^2,$$

where c_1 and c_2 are any complex numbers. If ψ_n denotes a discrete eigenfunction of α with the eigenvalue a_n, we have

$$f(\alpha)\psi_n = c_1\alpha\psi_n + c_2\alpha(\alpha\psi_n) = (c_1 a_n + c_2 a_n^2)\psi_n = f(a_n)\psi_n. \quad (37·4)$$

Thus ψ_n is an eigenfunction of $f(\alpha)$ as well as of α, provided that it conforms to the appropriate boundary conditions. In the most general case this means that ψ_n must belong to the adjoint manifold of $f(\alpha)$. If $f(\alpha)$ is real for every value of a_n, it means that ψ_n belongs to the Hermitian manifold of $f(\alpha)$.

Let us consider the most general case in which α is required to be a dynamical variable with an adjoint $\alpha^\dagger$. Let $T^{\binom{q}{z}}$ denote once more a unitary integral operator defining a canonical transformation which converts α to multiplicative form. Then

$$\alpha \equiv \alpha^{\binom{x}{z}} = (T^{-1})^{\binom{x}{q}}\alpha' T^{\binom{q}{z}}.$$

Equation (37·4) is seen to be equivalent to

$$f(\alpha) = (T^{-1})^{\binom{x}{q}}f(\alpha')T^{\binom{q}{z}} \tag{37·5}$$

or

$$f(\alpha)(x'|) = \lim_{\eta \to 0} \overline{\sum}_{q'}\phi_\eta(x'|q')f(\alpha')(\alpha'|).$$

Let $f(\alpha)^\dagger$ be defined by

$$f(\alpha)^\dagger \equiv (T^{-1})^{\binom{x}{q}}f(\alpha')^* T^{\binom{q}{z}}.$$

It follows [cf. Eq. (36·83)] that $f(\alpha)^\dagger$ is adjoint to $f(\alpha)$ with an adjoint manifold which includes all eigenfunctions and eigendifferentials of α together with all other functions $\psi(x)$ such that $\overline{\sum}_{q'}|f(\alpha')T^{\binom{q}{z}}\psi|^2$ converges. If $|f(\alpha')|$ has an upper bound, it is evident that the adjoint manifold of $f(\alpha)$ will contain all class D functions, so that $f(\alpha)$ generates a dynamical variable. The condition is sufficient but not necessary. In any case the eigenfunctions and eigendifferentials of $f(\alpha)$ are the same as those of α, while the eigenvalues are related by $f(\alpha)' = f(\alpha')$.

Equation (37·5) gives a scheme for defining functions of an operator with a complete orthonormal system of eigenfunctions which is alternative to the method of building up the function by the processes of multiplication and addition described in Sec. 37a. The same result is obtained if we adopt the method of von Neumann [cf. Eq. (36·35)] and write

$$f(\alpha)(x'|) = \int_{-\infty}^{+\infty}f(\sigma)d[E_\alpha(\sigma)(x'|)]. \tag{37·6}$$

If every eigenfunction of an operator α is also an eigenfunction of the operator β, it follows that all eigenfunctions of α with a common eigenvalue α' are also eigenfunctions of β with a common eigenvalue β'.[1]

---

[1] Otherwise we could choose a linear combination of eigenfunctions of α with a common eigenvalue which would not be an eigenfunction of β.

In that case each eigenvalue of β is a function of the corresponding eigenvalue of α and we say that β is a *function* of α. If two or more different eigenvalues of α correspond to the same eigenvalue of β, the relation is not a reciprocal one and α is not a function of β. If α is not a function of β and β is not a function of α, the operators α, β are said to be *independent*.

As has been emphasized by Dirac, functions of an operator defined by means of (37·4) and (37·5) can be just as discontinuous as functions of a real or complex number. Furthermore, this equation can be used to define the square root of an operator, or, in many cases, the reciprocal of one.[1]

If the operator α is Hermitian, it is easy to choose $f(\alpha')$ so that the function $f(\alpha)$ defined by (37·5) is unitary. For this purpose it is sufficient that $f(\alpha')$ shall map the axis of reals on the unit circle in the complex plane, so that $|f(\alpha')| = 1$ if α' is real (*cf.* p. 275).

A particularly important case is obtained by setting $f(x) = e^{i\lambda x}$, where λ is a real parameter. We denote the corresponding operator function of λ and α by $e^{i\lambda\alpha}$. Then

$$\frac{\partial}{\partial\lambda}[e^{i\lambda\alpha}(x'|)] = \frac{\partial}{\partial\lambda}[(T^{-1})^{\binom{x}{\beta}}e^{i\lambda\alpha'}T^{\binom{\beta}{x}}(x'|)] = (T^{-1})^{\binom{x}{\beta}}i\alpha'e^{i\lambda\alpha'}T^{\binom{\beta}{x}}(x'|)$$

$$= i[(T^{-1})^{\binom{x}{\beta}}\alpha'T^{\binom{\beta}{x}}][(T^{-1})^{\binom{x}{\beta}}e^{i\lambda\alpha'}T^{\binom{\beta}{x}}](x'|)$$

$$= i\alpha e^{i\lambda\alpha}(x'|).$$

Thus the function $\psi = e^{i\lambda\alpha}(x'|)$ is automatically a solution of the differential equation

$$\alpha\psi = \frac{1}{i}\frac{\partial\psi}{\partial\lambda},$$

which reduces to the form $(x'|)$ when λ is zero. If we set $\lambda = -\dfrac{2\pi t}{h}$ and identify α with the Hamiltonian operator H we secure in this way a solution of the second Schrödinger equation.

If the function $e^{i\lambda\alpha}(x'|)$ is analytic in λ, it is evident that by a repetition of the above differentiation process we can derive

$$e^{i\lambda\alpha}(x'|) = \sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\left\{\frac{\partial^n}{\partial\lambda^n}[e^{i\lambda\alpha}(x'|)]\right\}_{\lambda=0} = \sum_{n=1}^{\infty}\frac{(i\lambda\alpha)^n}{n!}(x'|).$$

Thus in this case the operator $e^{i\lambda\alpha}$ is equivalent to its formal power-series development (*cf.* Sec. 32*b*, p. 201).

37c. Commutative Operators.—In general the product of two operators depends upon the order of their application, *i.e.*, the commutative

---

[1] In the case of an operator α which has the discrete eigenvalue zero, the equation $\alpha^{-1}\psi_n = a_n^{-1}\psi_n$ has no meaning when a_n is zero. However, if 0 is in the continuous spectrum of α, α^{-1} can still exist.

law of ordinary multiplication does not apply.   Consider, for example, the basic operators

$$q_k \equiv [q_k \times], \quad p_l \equiv \frac{h}{2\pi i}\frac{\partial}{\partial q_l}.$$

In this case

$$(p_l q_k - q_k p_l)\psi = \frac{h}{2\pi i}\left[\frac{\partial}{\partial q_l}(q_k \psi) - q_k \frac{\partial\psi}{\partial q_l}\right] = \frac{h}{2\pi i}\delta_{kl}\psi, \qquad (37\cdot7)$$

or

$$p_l q_k - q_k p_l = \frac{h}{2\pi i}\delta_{kl}. \qquad (37\cdot8)$$

This "commutation" rule is characteristic of the behavior of so-called *conjugate* dynamical variables in quantum mechanics and will be the subject of further discussion in Sec. 39.

If α and β are any two operators, the operator $\dfrac{2\pi i}{h}(\beta\alpha - \alpha\beta)$ is called the *Poisson bracket*[1] of α and β and is indicated by the symbol $[\alpha,\beta]$.

$$[\alpha,\beta] \equiv \frac{2\pi i}{h}(\beta\alpha - \alpha\beta). \qquad (37\cdot9)$$

Thus Eq. (37·6) is usually written in the form

$$[q_k,p_l] = \delta_{kl}. \qquad (37\cdot10)$$

Definition: Two operators α and β are said to commute if their Poisson bracket yields zero when applied to any quadratically integrable functions for which it is defined.   When α and β do commute in this sense we can set $[\alpha,\beta]\psi$ equal to zero for every ψ whose transform by $[\alpha,\beta]$ is not defined *a priori*, and so insure that the transform of *every* quadratically integrable function by $[\alpha,\beta]$ shall exist and vanish identically.   We shall *assume* hereafter that α and β always commute when $[\alpha,\beta]\psi = 0$ for every ψ which is a member of some complete set of functions.   Exceptions to this rule can apparently occur but are deemed sufficiently unlikely to be disregarded for the purposes of quantum mechanics.

It follows from the theorem of p. 279 that if the representatives of two dynamical variables commute in any q' space, they commute in every legitimate coordinate space.   In that case we say that the dynamical variables themselves commute.

Any two multiplication operators in a given coordinate space, such as $[q_k'\times]$ and $[q_l'\times]$ must commute by the ordinary laws of algebra.   Furthermore, if the spectra of q_k and q_l are both purely continuous, as in

[1] The notation is taken over from classical dynamical theory (*cf.* Dirac, *P.Q.M.*, 2d ed., section 25).

the case of ordinary positional coordinates, the conjugate operators $p_k^{\binom{q}{q}} \equiv \dfrac{h}{2\pi i}\dfrac{\partial}{\partial q_k}$, and $p_l^{\binom{q}{q}} \equiv \dfrac{h}{2\pi i}\dfrac{\partial}{\partial q_l}$, commute by the elementary rules of partial differentiation. Finally (37·8) shows that $p_k^{\binom{q}{q}}$ and $q_l^{\binom{q}{q}}$ commute if q_k and q_l are independent variables with continuous spectra in a legitimate coordinate space.

If α is any dynamical variable and $f(\alpha)$ is any function of α, it follows that α commutes with $f(\alpha)$. This is a consequence of the fact that in a system of coordinates α_1', α_2', $\cdot\cdot\cdot$, of which α' is a member, the operators $\alpha^{\binom{\alpha}{\alpha}}$ and $f(\alpha)^{\binom{\alpha}{\alpha}}$ are multiplication operators.

A more interesting case of commutation is that of two (or more) independent dynamical variables α_1, α_2 whose representatives in Cartesian space $\alpha_1^{\binom{x}{x}}$ and $\alpha_2^{\binom{x}{x}}$ have a complete orthonormal system of simultaneous eigenfunctions. In the two-particle problem, for example, the three dynamical variables H, $\mathcal{L}^2$, $\mathcal{L}_z$ have simultaneous eigenfunctions

$$\psi_{nlm} = (2\pi)^{-\frac{1}{2}} R_{nl}(r)\Theta_{lm}(\theta)e^{im\varphi}.$$

Although no one of these operators is a function of either of the others, or conjugate to either of the others, they commute in pairs. To prove this we transform to a space in which the eigenvalues of the three operators are the coordinates. In this space all three are multiplication operators and must therefore commute. It follows that they commute also in the original Cartesian-coordinate space when applied to any quadratically integrable function. At the end of Sec. 38 it will be verified that the three operators H, $\mathcal{L}^2$, $\mathcal{L}_z$ commute in pairs with respect to all functions to which the product operators are applicable, whether the transforms are quadratically integrable or not.

Generalizing the above result, we see that if α_1, and α_2 are any two members of a mutually compatible set of dynamical variables whose eigenvalues form a legitimate coordinate system connected with the basic Cartesian system by means of a unitary transformation operator $T^{\binom{\alpha}{x}}$, it is *necessary* that α_1 and α_2 shall commute. In this case also $\alpha_1^{\binom{\alpha}{\alpha}}$ and $\alpha_2^{\binom{\alpha}{\alpha}}$ are multiplication operators which commute by the laws of ordinary algebra.

The question now arises whether the commutability of two independent dynamical variables α_1 and α_2 is also a *sufficient* condition for the existence of a coordinate system of which they are members. This question, or its equivalent in one form or another, is usually answered in the affirmative, but the proposition has not been rigorously established. The ingenious argument due to Dirac[1] and intended to cover the general

[1] P. A. M. Dirac, *P.Q.M.*, 1st ed., section 17.

case is unsatisfactory from the standpoint of rigor. An alternative discussion by von Neumann[1] is rigorous, but not so general in scope as might be desired. The following brief discussion is rigorous as far as it goes, but incomplete.

Let α and β denote the representatives in Cartesian-coordinate space of two commuting dynamical variables. Let D_α and D_β denote the class D adjoint manifolds of α and β, respectively. Let φ_{nm} be a normalized discrete eigenfunction of α with the eigenvalue a_n. We assume that a_n has a finite number, say g_n, of mutually orthogonal eigenfunctions so that the most general eigenfunction with the eigenvalue a_n is of the form $\sum_{m=1}^{g_n} \varphi_{nm} d_{nm}$. It follows from the commutation rule that

$$\alpha(\beta\varphi_{nm}) = \beta(\alpha\varphi_{nm}) = a_n(\beta\varphi_{nm}).$$

Let us now introduce the additional *assumption* that the transform of every φ_{nm} by β belongs both to D_α and to D_β. Then $\beta\varphi_{nm}$ vanishes identically, or is an eigenfunction of α. In the former case φ_{nm} is a simultaneous eigenfunction of β and α. In the latter case, if the eigenvalue a_n is nondegenerate ($g_n = 1$), $\beta\varphi_{nm}$ must be a multiple of φ_{nm}. Since it belongs to D_β, φ_{nm} is an eigenfunction of β as well as of α. If this is true of every eigenfunction of α, β is a function of α and the original φ_{nm}'s form a complete set of simultaneous eigenfunctions.

In the more general case where some of the a_n's are degenerate, $\beta\varphi_{nm}$ is necessarily a linear combination of the functions $\varphi_{n1}, \cdots, \varphi_{ng_n}$.

$$\beta\varphi_{nm} = \sum_{k=1}^{g_n} \varphi_{nk}\beta(nk;nm). \qquad (37\cdot11)$$

Here $\beta(nk;nm)$ is an expansion coefficient. The same type of expansion holds for every value of m from 1 to g_n. The coefficients $\beta(nk;nm)$ form a square matrix $\|\beta(nk;nm)\|$ with g_n rows and g_n columns. If the functions φ_{nm} of the original set happen to be all eigenfunctions of β as well as of α, the matrix will have the simple "diagonal" form

$$\|\beta(nk;nm)\| = \|\beta_{nk}\delta_{km}\|$$

in which all elements vanish except those on the principal diagonal (*cf.* Sec. 43, p. 349). If the matrix is not diagonal, we can replace the φ_{nm}'s by a new set of normalized orthogonal eigenfunctions of α, say $\psi_{n1}, \cdots, \psi_{ng_n}$ derived from the old by means of a linear transformation

$$\psi_{nm} = \sum_{k=1}^{g_n} \varphi_{nk}U_{km}.$$

It is then always possible to choose this transformation so that the new

[1] VON NEUMANN, *M.G.Q.*. II, 10, pp. 88–93.

functions form an orthonormal set of simultaneous eigenfunctions of both operators. The proof of this proposition is algebraic and will be supplied in connection with our general examination of the matrix form of the eigenvalue problem in Sec. 44c. If α has a discrete spectrum only, so that the φ's form a complete set, the ψ's will form a similar complete set.

Fortunately the hypotheses on which the above proof of the existence of a complete system of simultaneous eigenfunctions of α and β is based can be somewhat relaxed without loss in rigor. The technique involved in extending the theorem to the weaker hypotheses (due to von Neumann) is too elaborate for reproduction here and we content ourselves with a statement of the result. Let α and β denote two Hermitian[1] operators for which the eigenvalue-eigenfunction problem is solvable. Let M_a denote the linear manifold of eigenfunctions of α for the arbitrary eigenvalue a. *It suffices for the existence of a complete set of simultaneous eigenfunctions of α and β that* (**a**) *they commute,* (**b**) *one of the operators, say α, has a purely discrete spectrum;* (**c**) *independent of the value of a it is possible to approximate every function ψ in M_a with any desired precision by an element of M_a which belongs to the Hermitian manifold of β.* von Neumann does not give this theorem explicitly, but it is implicit in the discussion of commuting operators which he does give (*M.G.Q.*, II, 10). A generalization of the theorem to cases in which neither operator has a purely discrete spectrum is much to be desired but presents serious mathematical difficulties.

It follows as a corollary on the above theorem that if α does not have a purely discrete spectrum but does have the discrete eigenvalue a, every function in M_a can be expanded into a series (or integral) of simultaneous eigenfunctions of α and β. To verify the corollary one needs only to observe that it is always possible to define an operator γ with two eigenvalues 0 and 1, such that every function in M_a belongs to the eigenvalue 1 and every quadratically integrable function orthogonal to the elements of M_a belongs to the eigenvalue 0. This operator is then a function of α which satisfies the above postulates (**a**), (**b**), (**c**). Since γ and β have a complete set of simultaneous eigenfunctions, every function in M_a can be expanded in terms of them. This proves the proposition.

Let us now identify α with the Hamiltonian operator H, and β with an arbitrary dynamical variable which commutes with H. The discrete eigenfunctions of H belong to class D and the adjoint manifold of β includes class D. Hence (**c**) is satisfied for every discrete energy level $E_k = a$ and it follows from the above corollary that *every eigenfunction*

---

[1] Since every dynamical-variable operator gives rise to a Hermitian operator with the same set of eigenfunctions, and *vice versa*, the theorem must hold if α and β are dynamical variables which are not necessarily Hermitian, provided that we substitute the phrase "adjoint manifold of β" for "Hermitian manifold of β."

of H which belongs to E_k can be expanded into a series (or integral) of simultaneous eigenfunctions of H and β.

Experience goes beyond our proofs to show that *ordinarily* any two dynamical variables which commute satisfy the condition (c) and have a complete system of simultaneous eigenfunctions. Hence we shall designate any set of dynamical variables which have such a system of eigenfunctions as a *normal commuting set*. The variables which form such a set are *mutually compatible* in the sense of Sec. 36d. There exists then a legitimate coordinate system of which any two normal commuting dynamical variables α_1, α_2 are members, with a unitary operator $T^{\binom{\alpha}{x}}$ which transforms $\alpha_1 = \alpha_1^{\binom{x}{x}}$ and $\alpha_2 = \alpha_2^{\binom{x}{x}}$ simultaneously to multiplicative form.

37d. Functions of a Normal Set of Commuting Dynamical Variables. A function of such a set of commuting dynamical variables $\alpha_1, \alpha_2, \cdots, \alpha_\lambda$ is defined as a dynamical variable whose representative in α' space has the form

$$f(\alpha_1, \alpha_2, \cdots, \alpha_\lambda)^{\binom{\alpha}{x}} = [f(\alpha_1', \cdots, \alpha_\lambda') \times].$$

In other words, it is an operator f such that every simultaneous eigenfunction of the members of the commuting set is also an eigenfunction of f. It follows from this definition that if we can prepare an assemblage of physical systems with a unique set of values of $\alpha_1, \alpha_2, \cdots, \alpha_\lambda$ it must necessarily have a unique value of f. If no one member of a set of normal commuting dynamical variables is a function of the others, the variables are said to be mutually *independent*.

Following the argument of Sec. 36d, p. 250, it is easy to show that if some of the sets of simultaneous discrete eigenvalues of a normal group of commuting dynamical variables exhibit a finite degeneracy, it is always possible to remove that degeneracy by the addition of one or more independent commuting dynamical variables to the group.[1] Extending an hypothesis introduced in Sec. 36d, we postulate that all the degeneracy of the simultaneous eigenvalues of such a group, for both discrete and continuous spectra, can always be removed in this way. On the other hand, if none of the simultaneous eigenvalues of such a set of normally commuting operators, say $\alpha_1, \alpha_2, \cdots, \alpha_\lambda$, is degenerate, any operator β which commutes normally with all members of the group must be a function of the α's. This follows from the fact that there exists a complete system of simultaneous eigenfunctions of $\alpha_1, \cdots, \alpha_\lambda, \beta$, each of which is determined to a constant factor by the eigenvalues of the α's alone. Consequently every simultaneous eigenfunction of the α's must be *a priori* an eigenfunction of β, and every set of eigenvalues of

[1] *Cf.* P. A. M. Dirac, *P.Q.M.*, 1st ed., pp. 45–46.

the α's must determine a unique corresponding eigenvalue of β. Hence a set of normally commuting dynamical variables is said to be *complete* if every one of its sets of simultaneous eigenvalues is nondegenerate. Clearly we can with advantage restrict ourselves to the use of coordinate systems based on *complete sets of normally commuting independent dynamical variables*. The number of operators in such a complete set is *not* a characteristic of the dynamical system under investigation—like the number of degrees of freedom in classical mechanics—but depends on the choice of operators. In fact we can in principle replace any complete set of normally commuting independent dynamical variables by a single dynamical variable without changing the system of eigenfunctions. To this end it is only necessary to pass a single line through all allowed points of α' space and to correlate points on this line with distance from its starting point. This is possible in principle by a well-known mathematical theorem.

Theorem:[1] *If the dynamical variables α and β form a normal commuting pair, α commutes with any function of β and β with any function of α.* The proof is an immediate consequence of the fact that α, β, and all functions of α, or of β, are represented by multiplication operators in any coordinate system of which α' and β' are elements.

Theorem: If every dynamical variable which commutes with the dynamical variable α commutes also with a second dynamical variable β, it follows that β is a function of α.

Let γ be an arbitrary dynamical variable which commutes with α and hence with β. We shall consider the representatives of α, β, and γ in a space $\alpha_1', \alpha_2', \cdots$, in which one coordinate, say α_1', gives the eigenvalues of α. The theorem then reduces to the statement that, under the conditions stated, $\beta^{(\alpha)}_{(\alpha)}$ has the form $[w(\alpha_1')\times]$. Let $\gamma^{(\alpha)}_{(\alpha)}$ be any multiplication operator, say $[\varphi(\alpha_1',\alpha_2', \cdots)\times]$ and let $u(\alpha')$ denote an arbitrary probability amplitude. As $\gamma^{(\alpha)}_{(\alpha)}$ is a multiplication operator, it commutes with $\alpha^{(\alpha)}_{(\alpha)}$ and hence with $\beta^{(\alpha)}_{(\alpha)}$ as well. Thus $\beta^{(\alpha)}_{(\alpha)}\varphi u = \varphi\beta^{(\alpha)}_{(\alpha)}u$. It follows that for arbitrary choices of the functions u and φ and at points where neither vanishes

$$\frac{\beta^{(\alpha)}_{(\alpha)}(\varphi u)}{\varphi u} = \frac{\beta^{(\alpha)}_{(\alpha)}u}{u} = w(\alpha_1',\alpha_2', \cdots).$$

This is sufficient to establish the multiplicative character of the operator $\beta^{(\alpha)}_{(\alpha)}$. It remains to show that the function $w(\alpha')$ is actually independent of all the α''s except α_1'. This can be done if we can find a set of non-multiplicative operators in α' space which commute with the operation

[1] *Cf.* DIRAC, *op. cit.*, p. 41.

of multiplying by α_1' but not with any of the other operators $[\alpha_2'\times]$, $[\alpha_3'\times]$, $\cdots$. If α_k, for example, has a purely continuous spectrum ranging from $-\infty$ to $+\infty$, the operator $i\partial/\partial\alpha_k'$ represents a dynamical variable in α' space, commutes with $[\alpha_1'\times]$ but not with $[\alpha_k'\times]$. By hypothesis this operator must commute with $\beta^{\left(\frac{\alpha}{\alpha}\right)}$, which means that

$$\frac{\partial}{\partial\alpha_k'}wu = w\frac{\partial}{\partial\alpha_k'}u$$

for every differentiable u. Thus $\partial w/\partial\alpha_k'$ is identically equal to zero, and w is independent of α_k'. The more general case where the spectrum of α_k is partially continuous and partially discrete can be dealt with by means of suitably defined exchange operators. It will suffice to consider two points on the discrete spectrum of α_k which we label α_k' and α_k''. To prove that w has the same value at any pair of corresponding points $P_1 = (\alpha_1',\alpha_2', \cdots ,\alpha_k', \cdots)$, $P_2 = (\alpha_1',\alpha_2', \cdots ,\alpha_k'', \cdots)$ we identify $\gamma^{\left(\frac{\alpha}{\alpha}\right)}$ with the operator that interchanges the values of $u(\alpha')$ at every such pair of points, while leaving $u(\alpha')$ unchanged everywhere else. The operator is readily seen to fulfill the requirements for a dynamical variable and commutes with $[\alpha_1'\times]$. It can commute with $[w\times]$, however, only if w has the same value at every pair of corresponding points. As the values of α_k' and α_k'' are arbitrary, it follows that w is independent of α_k' over the entire discrete spectrum. Similarly we show that it is independent of α_k' over the entire continuous spectrum. To demonstrate that w has the same value for a pair of corresponding points P_1, P_2 in this spectrum, a more general type of Hermitian operator $\gamma^{\left(\frac{\alpha}{\alpha}\right)}$ is needed; we omit the discussion of this final step in the proof.

Obviously we can replace the single dynamical variable α by a normal set of commuting dynamical variables in each of the foregoing theorems.

38. THE CONSERVATION LAWS

38a. Conservation of Energy.—In the classical mechanics the energy and the Hamiltonian function $H(p,q)$ of a conservative system are constant in time for any natural motion. In quantum mechanics we have, in general, no single definite energy but a distribution function giving the probability of various energy levels. By "conservation of energy" we may therefore imply the constancy of this distribution function in time, or, what comes to the same thing, the constancy of the mean values of H and its various powers. If the Hamiltonian operator $H\left(\frac{\partial}{\partial q},q\right)$ does not depend upon the time explicitly—as hitherto assumed— conservation is readily proved with the aid of the expansion (35·1) applied to a class D solution of the second Schrödinger equation (7·6). It follows directly from the method of calculating energy probabilities described in

Sec. 35a [cf. especially Eq. (35·2)] that the distribution function which gives the probability of different energy values is independent of t.

38b. Variation of Energy When the Hamiltonian Depends on the Time.—In the case of a system for which the classical Hamiltonian function and the corresponding operator depend explicitly on the time, the energy is not conserved classically and the eigenfunctions of the energy operator do not yield monochromatic solutions of the second Schrödinger equation. It is still possible, however, to define an energy distribution by expanding Ψ into a linear combination of eigenfunctions of the equation

$$H\left(\frac{\partial}{\partial q}, q, t\right)\Psi = E\Psi, \qquad (38\cdot1)$$

in which t is treated as a parameter. This amounts to assigning to E at each instant the energy spectrum and distribution function which would be appropriate if the operator H could be frozen so that its instantaneous form became permanent. The coefficients in this expansion will no longer be simple exponential functions of the time, however, and hence the distribution function defined by (36·68) and (36·77) will vary with the time. The average energy will be given by

$$\bar{E} = \int_{\infty} \Psi^* H\left(\frac{\partial}{\partial q}, q, t\right)\Psi d\tau \qquad (38\cdot2)$$

and is obviously dependent on t. Taking into account the Hermitian character of H we readily derive

$$\frac{d\bar{E}}{dt} = \int_{\infty}\left[\frac{\partial \Psi^*}{\partial t}H\Psi + \Psi^*\frac{\partial(H\Psi)}{dt}\right]d\tau = \int_{\infty}\Psi^*\left[\frac{\partial(H\Psi)}{\partial t} - H\frac{\partial\Psi}{\partial t}\right]d\tau.$$

In the case of an atom or molecule, subject to the influence of an external classical radiation field (cf. Sec. 7), we can write the Hamiltonian as the sum of two terms, one of which represents the energy in the absence of the field while the other varies with t and gives the contribution of the field to the motion. The second, or mutual, energy term is ordinarily assumed to be small. To measure the nonmutual energy at any instant t it would be necessary to isolate the system by removing the radiation field and then to determine the energy by a method appropriate to the isolated or unperturbed system. This experimental procedure corresponds in theory to the use of an energy spectrum and distribution function obtained by analyzing the instantaneous Ψ function into a linear combination of eigenfunctions of the equation

$$H_0\left(\frac{\partial}{\partial q}, q\right)\psi = E\psi, \qquad (38\cdot3)$$

where H_0 is the constant part of the Hamiltonian operator.[1] The variation in the distribution function thus defined (*cf.* Sec. 54) is commonly interpreted as due to "quantum jumps" from one energy level to the other caused by the radiation.

38c. Conservation of an Arbitrary Dynamical Variable.—Consider next the conservation problem for an arbitrary dynamical variable α. We assume that the energy itself is conserved and that α does not depend explicitly upon the time so that it commutes with the operator $\partial/\partial t$. *Under these conditions a sufficient condition for the conservation of α is that H and α form a normal commuting pair of dynamical variables.*

To prove this theorem let $T^{\binom{\beta}{x}}$ denote the unitary operator that transforms probability amplitudes in Cartesian coordinate space into probability amplitudes based on a coordinate system β_1, β_2, $\cdots$ of which α is a member. The distribution function giving the probability of different eigenvalues of α for the Cartesian wave function $\Psi(x,t)$ is

$$Q(\alpha') = \overline{\sum}'_{\beta'} |T^{\binom{\beta}{x}}\Psi(x,t)|^2, \tag{38.4}$$

where $\overline{\sum}'_{\beta'}$ denotes a sum integration over all the β' coordinates except α'. As H and α form a normal commuting pair, we can include H among the dynamical variables β_1, β_2, $\cdots$. Then, since $\Psi(x,t)$ is a solution of the second Schrödinger equation, it follows from Sec. 37b that

$$\Psi(x,t) = (T^{-1})^{\binom{x}{\beta}} e^{-\frac{2\pi i Et}{h}} T^{\binom{\beta}{x}}\Psi(x,0).$$

Consequently

$$\frac{dQ(\alpha')}{dt} = \frac{d}{dt}\overline{\sum}'_{\beta'} |e^{-\frac{2\pi i Et}{h}} T^{\binom{\beta}{x}}\Psi(x,0)|^2 = \frac{d}{dt}\overline{\sum}'_{\beta'} |T^{\binom{\beta}{x}}\Psi(x,0)|^2 \equiv 0, \tag{38.5}$$

as was to be proved.

An interesting alternative proof is obtained if we note that in case the distribution function for α is not constant in the neighborhood of $\alpha' = \alpha''$ it must be possible to choose a function $f(\alpha')$ which vanishes outside the immediate neighborhood of $\alpha' = \alpha''$ and such that $\frac{d\overline{f(\alpha')}}{dt}$ does not vanish. Since α and H are assumed to form a normal commuting pair, $f(\alpha)\Psi$ will belong to the Hermitian domain of H and we have

$$\frac{d\overline{f(\alpha)}}{dt} = \int_\infty \left[f\Psi\frac{\partial\Psi^*}{\partial t} + \Psi^* f\frac{\partial\Psi}{\partial t}\right]d\tau = \frac{2\pi i}{h}\int_\infty [H^*\Psi^* f\Psi - \Psi^* fH\Psi]d\tau$$

$$= \frac{2\pi i}{h}\int_\infty [\Psi^*(Hf - fH)\Psi d\tau \doteq 0. \tag{38.6}$$

Thus we get a contradiction which proves the theorem.

[1] *Cf.* KEMBLE and HILL, *Rev. Mod. Phys.* **2**, 6 (1930).

In classical mechanics dynamical variables which remain constant during the natural motion of the system are called *integrals of the equations of motion* or, simply, *integrals of the motion*. The analogue of such a classical "integral" in quantum mechanics is the dynamical variable which unites with the Hamiltonian H to form a normal commuting pair and whose distribution function therefore remains constant in time when computed from a solution of the Schrödinger equation $H\Psi = -\dfrac{h}{2\pi i}\dfrac{\partial\Psi}{\partial t}$. We shall therefore refer to such dynamical variables as *integrals of the motion* or *integrals of the Schrödinger equation*. Since an arbitrary solution of the Schrödinger equation for the hydrogenic-atom problem can be developed in terms of simultaneous eigenfunctions of H, $\mathcal{L}^2$, and $\mathcal{L}_x$, $\mathcal{L}_y$, or $\mathcal{L}_z$, each of these operators is an integral of the motion for such a system. In Sec. 38d it will be proved that these operators are integrals of the motion for any free atomic or molecular system.

38d. Commutation Properties of the Hamiltonian and the Angular Momentum.—Let us now briefly consider the commutation properties of the five operators just mentioned in the general case of an n-particle problem, starting from the definitions

$$\vec{r}_k = \vec{i}\,x_k + \vec{j}\,y_k + \vec{k}\,z_k, \tag{38.7}$$

$$\vec{p}_k = \vec{i}\,\xi_k + \vec{j}\,\eta_k + \vec{k}\,\zeta_k = \frac{h}{2\pi i}\left(\vec{i}\frac{\partial}{\partial x_k} + \vec{j}\frac{\partial}{\partial y_k} + \vec{k}\frac{\partial}{\partial z_k}\right) = \frac{h}{2\pi i}\overrightarrow{\mathrm{grad}}_k, \tag{38.8}$$

$$H = \sum_{k=1}^{n}\frac{[\xi_k^2 + \eta_k^2 + \zeta_k^2]}{2\mu_k} + V(x_1,\cdots,z_n) = -\frac{h^2}{8\pi^2}\sum_k\frac{1}{\mu_k}\nabla_k^2 + V, \tag{38.9}$$

$$\mathcal{L}_x = \sum_k\mathcal{L}_{kx} = \sum_k(y_k\zeta_k - z_k\eta_k) = \frac{h}{2\pi i}\sum_k\left(y_k\frac{\partial}{\partial z_k} - z_k\frac{\partial}{\partial y_k}\right)$$

$$= \frac{h}{2\pi i}\sum_k[\vec{r}_k \times \overrightarrow{\mathrm{grad}}_k]_x, \tag{38.10}$$

$$\mathcal{L}_y = \sum_k\mathcal{L}_{ky} = \sum_k(z_k\xi_k - x_k\zeta_k) = \frac{h}{2\pi i}\sum_k\left(z_k\frac{\partial}{\partial x_k} - x_k\frac{\partial}{\partial z_k}\right)$$

$$= \frac{h}{2\pi i}\sum_k[\vec{r}_k \times \overrightarrow{\mathrm{grad}}_k]_y, \tag{38.11}$$

$$\mathcal{L}_z = \sum_k\mathcal{L}_{kz} = \sum_k(x_k\eta_k - y_k\xi_k) = \frac{h}{2\pi i}\sum_k\left(x_k\frac{\partial}{\partial y_k} - y_k\frac{\partial}{\partial x_k}\right)$$

$$= \frac{h}{2\pi i}\sum_k[\vec{r}_k \times \overrightarrow{\mathrm{grad}}_k]_z, \tag{38.12}$$

$$\mathcal{L}^2 = \mathcal{L}_x^2 + \mathcal{L}_y^2 + \mathcal{L}_z^2. \tag{38.13}$$

We note first of all that each of the elementary operators $x_1, \cdots, z_n,$ $\xi_1, \cdots, \zeta_n$ commutes with every member of the set except its own conjugate. Then, by (38·10),

$$(\xi_k{}^2 + \eta_k{}^2 + \zeta_k{}^2)\mathcal{L}_x - \mathcal{L}_x(\xi_k{}^2 + \eta_k{}^2 + \zeta_k{}^2)$$
$$= (\eta_k{}^2 + \zeta_k{}^2)(y_k\zeta_k - z_k\eta_k) - (y_k\zeta_k - z_k\eta_k)(\eta_k{}^2 + \zeta_k{}^2)$$
$$= \zeta_k(\eta_k{}^2 y_k - y_k\eta_k{}^2) - \eta_k(\zeta_k{}^2 z_k - z_k\zeta_k{}^2)$$
$$= \zeta_k\left(\eta_k y_k \eta_k + \frac{h}{2\pi i}\eta_k - y_k\eta_k{}^2\right) - \eta_k\left(\zeta_k z_k\zeta_k + \frac{h}{2\pi i}\zeta_k - z_k\zeta_k{}^2\right)$$
$$= \zeta_k\left(\frac{h}{\pi i}\eta_k\right) - \eta_k\left(\frac{h}{\pi i}\zeta_k\right) = 0.$$

Hence

$$[\mathcal{L}_x, H] = [\mathcal{L}_x, V] = -\sum_k \{[\vec{r}_k \times \overrightarrow{\mathrm{grad}}_k]_x V - V[\vec{r}_k \times \overrightarrow{\mathrm{grad}}_k]_x\}$$
$$= -\sum_k [\vec{r}_k \times (\overrightarrow{\mathrm{grad}}_k\, V)]_x. \tag{38·14}$$

The right-hand member of the above equation is an operator which stands for multiplication by a function of the coordinates, viz., by $h/2\pi i$ times the x component of the classical torque applied to the system in the configuration $x_1, \cdots, z_n$. The Poisson brackets of H with $\mathcal{L}_y$ and $\mathcal{L}_z$ are obtained by permuting the letters cyclically. If the external forces applied to the system have spherical symmetry, the classical applied torque will always be zero and $\mathcal{L}_x, \mathcal{L}_y, \mathcal{L}_z$ will commute with H. This statement applies in particular to a group of electrons moving about a fixed, positively charged nucleus. If the force field has axial symmetry, as in the case of an atom subject to a uniform electric or magnetic field, the component of angular momentum parallel to the axis of symmetry commutes with H, but the other components do not. If any of these operators commutes with a Hamiltonian H the sufficient conditions for a normal commuting pair specified on p. 285 are clearly fulfilled if we identify α with the operator in question and β with H. Hence each of these operators is an integral of the motion when it commutes with H.

Clearly, if $\mathcal{L}_x, \mathcal{L}_y, \mathcal{L}_z$ all commute with H, $\mathcal{L}^2$ must also commute with H. Hence $\mathcal{L}^2$ is an integral for systems of spherical symmetry, and for such cases only.

Finally we note that

$$(y_k\zeta_k - z_k\eta_k)(z_k\xi_k - x_k\zeta_k) - (z_k\xi_k - x_k\zeta_k)(y_k\zeta_k - z_k\eta_k)$$
$$= y_k\xi_k(\zeta_k z_k - z_k\zeta_k) - x_k\eta_k(\zeta_k z_k - z_k\zeta_k) = \frac{h}{2\pi i}(y_k\xi_k - x_k\eta_k).$$

Hence

$$[\mathcal{L}_x, \mathcal{L}_y] = \mathcal{L}_z. \tag{38·15}$$

Advancing the letters cyclically we obtain

$$[\mathcal{L}_y, \mathcal{L}_z] = \mathcal{L}_x, \tag{38·16}$$

$$[\mathcal{L}_z, \mathcal{L}_x] = \mathcal{L}_y. \tag{38·17}$$

These commutation relations can be thrown into the vector form

$$\vec{\mathcal{L}} \times \vec{\mathcal{L}} = -\frac{h}{2\pi i}\vec{\mathcal{L}}, \tag{38·18}$$

if due attention is paid to the order of the factors in the various terms of each component of the vector product $\vec{\mathcal{L}} \times \vec{\mathcal{L}}$. The operator $\vec{\mathcal{L}}_k$ can be written in the form $\frac{h}{2\pi i}\vec{r}_k \times \overrightarrow{\text{grad}}_k$ and transforms like a three-dimensional vector. Hence both sides of (38·18) transform like vectors and make this equation invariant of a rotation of the axes.

39. CONJUGATE DYNAMICAL VARIABLES AND QUANTUM-MECHANICAL EQUATIONS OF MOTION

39a. Conjugate Dynamical Variables.—In classical theory the component of momentum canonically conjugate to the generalized positional coordinate q_k is the quantity

$$p_k = \frac{\partial}{\partial \dot{q}_k} L(q, \dot{q}), \tag{39·1}$$

where L is the Lagrangian function of the coordinates and their velocities (*cf.* footnote 1, p. 23). If the forces are derivable from a potential function the components of linear momentum canonically conjugate to the Cartesian coordinates x_k, y_k, z_k are

$$\xi_k = \mu_k \dot{x}_k, \qquad \eta_k = \mu_k \dot{y}_k, \qquad \zeta_k = \mu_k \dot{z}_k,$$

respectively. In the more general case of an electrified particle moving in an external electromagnetic field (*cf.* Sec. 7d) the components of linear momentum defined by (39·1) have the form indicated by the vector equation (15·24).

The introduction of the momenta p_k along with the coordinates q_k permits the reduction of the classical equations of motion to the first-order canonical Hamiltonian form

$$\frac{dq_k}{dt} = \frac{\partial H(p, q, t)}{\partial p_k}, \qquad \frac{dp_k}{dt} = -\frac{\partial H(p, q, t)}{\partial q_k}. \tag{39·2}$$

This form is valid, however, not only for an arbitrary set of positional coordinates q_k with their corresponding momenta but also if we express the Hamiltonian function in terms of any set of variables

$$Q_k = Q_k(q_k, p_k, t), \qquad P_k = P_k(q_k, p_k, t)$$

of a wider class derived from any initial set of p's and q's by means of a contact transformation which involves both the coordinates and the momenta. Hence the term "canonically conjugate" is applied in classical mechanics to any pair of variables Q_k, P_k of a double set which permits the reduction of the equations of motion to the form

$$\frac{dQ_k}{dt} = \frac{\partial H}{\partial P_k}, \qquad \frac{dP_k}{dt} = -\frac{\partial H}{\partial Q_k}.$$

In defining conjugate dynamical variables in quantum mechanics one can seek pairs of operators associated with known pairs of conjugate classical variables, or one can make direct use of the quantum-mechanical analogue of the Hamiltonian equations. Starting from the former standpoint we let x_k and p_k denote, respectively, a classical Cartesian coordinate and its conjugate momentum. The corresponding dynamical variables in quantum mechanics are defined by the representative operators

$$x_k^{(x)} = [x_k' \times], \qquad p_{x_k}^{(x)} = \frac{h}{2\pi i} \frac{\partial}{\partial x_k'}. \qquad (39\cdot3)$$

As another example, consider the z component of angular momentum which is classically conjugate to the azimuthal coordinate $\alpha_1 = \varphi_1$ in a system of coordinates r_k, θ_k, α_k derived from an ordinary spherical system r_k, θ_k, φ_k by the transformation

$$\alpha_1 = \varphi_1, \qquad \alpha_k = \varphi_k - \varphi_1. \qquad k = 2, 3, \cdots.$$

We have already identified the quantum variable $\mathcal{L}_z$ with the class of operators whose representative in Cartesian coordinates is

$$\mathcal{L}_z^{(x)} = \sum_k \frac{h}{2\pi i}\left(x_k' \frac{\partial}{\partial y_k'} - y_k' \frac{\partial}{\partial x_k'}\right).$$

Furthermore it was proved in Sec. 34d that a "direct transformation" of $\mathcal{L}_z^{(x)}$ to the r_k, θ_k, α_k system of coordinates yields the operator $\frac{h}{2\pi i} \frac{\partial}{\partial \alpha_1}$. The direct transformation, however, did not allow for a renormalization of the probability amplitudes in the new coordinate system and we accordingly make use of the basic formula (36·5) to get the correct representative of $\mathcal{L}_z$ in the r_k, θ_k, α_k system, or the q' system, as we shall hereafter call it.

$$\mathcal{L}_z^{(q)} = T^{(q)} \mathcal{L}_z^{(x)} (T^{-1})^{(x)}_{(q)}. \qquad (39\ 4)$$

Here $T^{(q)}$ denotes the operator which transforms a probability amplitude $(x'|)$ into the corresponding probability amplitude $(q'|)$. Since this is a transformation from one system of positional coordinates to another

(*cf.* p. 239), $T^{\left(\begin{smallmatrix} q \\ z \end{smallmatrix}\right)}$ reduces to the product of the square root of the Jacobian

$$D(q') = \left| \frac{\partial(r_1', \theta_1', \cdots, \alpha_f')}{\partial(x_1', y_1', \cdots, z_f')} \right|$$

into the operator $\Theta^{\left(\begin{smallmatrix} q \\ x \end{smallmatrix}\right)}$ which denotes the substitution for each Cartesian coordinate of its value in terms of the q' coordinates. Similarly $(T^{-1})^{\left(\begin{smallmatrix} x \\ q \end{smallmatrix}\right)}$ can be written in the form $\Theta^{\left(\begin{smallmatrix} x \\ q \end{smallmatrix}\right)} D(q')^{-\frac{1}{2}}$. Thus

$$\mathcal{L}_z{}^{\left(\begin{smallmatrix} q \\ q \end{smallmatrix}\right)} = D(q')^{\frac{1}{2}}\Theta^{\left(\begin{smallmatrix} q \\ x \end{smallmatrix}\right)}\mathcal{L}_z{}^{\left(\begin{smallmatrix} x \\ x \end{smallmatrix}\right)}\Theta^{\left(\begin{smallmatrix} x \\ q \end{smallmatrix}\right)} D(q')^{-\frac{1}{2}}. \tag{39.5}$$

The operator $\Theta^{\left(\begin{smallmatrix} q \\ x \end{smallmatrix}\right)}\mathcal{L}_z{}^{\left(\begin{smallmatrix} x \\ x \end{smallmatrix}\right)}\Theta^{\left(\begin{smallmatrix} x \\ q \end{smallmatrix}\right)}$ is the *direct transform* of $\mathcal{L}_z{}^{\left(\begin{smallmatrix} x \\ x \end{smallmatrix}\right)}$ in the q' coordinate system as derived by the ordinary methods of partial differentiation. The Jacobian $D(q')$ turns out to be the product $\prod_k r_k'^2 \sin \theta_k'$. Thus, by (34·13),

$$\mathcal{L}_z{}^{\left(\begin{smallmatrix} q \\ q \end{smallmatrix}\right)} = \left[\prod_k r_k'^2 \sin \theta_k' \right]^{\frac{1}{2}} \frac{h}{2\pi i} \frac{\partial}{\partial \alpha_1'} \left[\prod_k r_k'^2 \sin \theta_k' \right]^{-\frac{1}{2}} = \frac{h}{2\pi i} \frac{\partial}{\partial \alpha_1'}. \tag{39.6}$$

The parallelism between (39·3) and (39·6) suggests that in all cases where q_k is a positional coordinate we can correlate quantum-mechanical conjugates with classical conjugates by defining the momentum quantum-mechanically conjugate to q_k by the formula

$$p_k{}^{\left(\begin{smallmatrix} q \\ q \end{smallmatrix}\right)} = \frac{h}{2\pi i} \frac{\partial}{\partial q_k'}. \tag{39.7}$$

By so doing we maintain the commutation rule (37·8) for all such pairs of dynamical variables.

As a check on this procedure we must verify, if possible, that the operator $p_k{}^{\left(\begin{smallmatrix} q \\ q \end{smallmatrix}\right)}$ so defined will always satisfy the restrictions imposed on dynamical variables. First of all comes the question of its Hermitian character.

Let the spectrum of q_k extend from $q_k' = a$ to $q_k' = b$. Clearly

$$\left(\frac{h}{2\pi i} \frac{\partial \psi_q^{(1)}}{\partial q_k'}, \psi_q^{(2)} \right) - \left(\psi_q^{(1)}, \frac{h}{2\pi i} \frac{\partial \psi_q^{(2)}}{\partial q_k'} \right)$$

$$= \frac{h}{2\pi i} \int \cdots \int [\psi_q^{(1)}\psi_q^{(2)*}] \Big|_a^b dq_1' \cdots dq_{k-1}'dq_{k+1}' \cdots dq_f'. \tag{39.8}$$

It follows that the operator $\dfrac{h}{2\pi i} \dfrac{\partial}{\partial q_k'}$ is Hermitian with respect to the class of physically admissible functions ψ_q provided that the scalar products on the left are convergent for every pair of functions $\psi_q^{(1)}$, $\psi_q^{(2)}$ belonging

to this class and provided that every such function takes on the same values at the end points a and b of the spectrum of q_k.

Consider the domains

$$q_k'(x_1', \cdots, x_f') = a, \qquad q_k'(x_1', \cdots, x_f') = b$$

in Cartesian or X space. Since *all* X space is mapped on that portion of q space between $q_k' = a$ and $q_k' = b$, these boundary domains of q space must either map coincident hypersurfaces in X space or else they must in some sense map boundary domains in Cartesian space.

Consider, for example, the spherical coordinates r, θ, φ. The angle φ ranges from zero to 2π and the surfaces $\varphi = 0$ and $\varphi = 2\pi$ are coincident half planes in X space. The angle θ ranges from zero to π and the domains $\theta = 0$ and $\theta = \pi$ of r,θ,φ space map the two halves of the z axis in X space. As the z axis is a line and not a surface, we can regard it as a boundary in X space without thereby removing any volume. The radius r ranges from zero to infinity, its boundary values marking spheres of zero and infinite radius, respectively, in X space.

If the domains $q_k' = a$, $q_k' = b$ are coincident hypersurfaces in X space, $\psi_q|_a^b = D^{\frac{1}{2}}\psi_x|_a^b$ vanishes automatically because ψ_x is single-valued. If these domains are not coincident hypersurfaces, ψ_q will ordinarily vanish on both of them. There are two cases to be considered. The first is that in which $q_k' = a$ corresponds to infinite values of one or more of the Cartesian coordinates. If ψ_x is a class D function vanishing exponentially at infinity, we can assume without appreciable loss of generality that $D^{\frac{1}{2}}\psi_x$ vanishes at infinity or that $\lim_{q_k' \to a} \psi_q = 0$. The second case is that, typified by the limits $\theta = 0$ and $\theta = \pi$ above, in which the domain $q_k' = a$ maps a degenerate domain in X space having dimensions less than $f - 1$, f being the number of dimensions in X space. Under these circumstances an infinitesimal volume of X space in the neighborhood of the domain in question will be mapped on a relatively very large volume of q space, so that the Jacobian D must vanish at $q_k' = a$.

Thus the right-hand member of (39·8) will practically always vanish The integrals on the left will usually converge and we conclude that if q_k is a positional coordinate obtained from the Cartesian coordinates by an ordinary point transformation, $\dfrac{h}{2\pi i} \dfrac{\partial}{\partial q_k}$, is normally Hermitian with respect to physically admissible functions in q space, *i.e.*, with respect to the linear manifold of functions $D^{(q)}$ obtained by applying the operator $T^{(q)}$ to the functions of class D.

It does not follow, however, that $\dfrac{h}{2\pi i} \dfrac{\partial}{\partial q_k}$ is the representative in the q coordinate system of a satisfactory real dynamical variable p_k. For an example of a classical positional coordinate which has no satisfactory

quantum-mechanical conjugate we need go no farther than the radius r in a three-dimensional spherical coordinate system. The operator $\frac{h}{2\pi i}\frac{\partial}{\partial r'}$ is easily seen to be Hermitian with respect to physically admissible functions of r', θ', φ' as these functions all vanish at the origin and at infinity. Solutions of the equation

$$\frac{h}{2\pi i}\frac{\partial \psi}{\partial r'} = p_r'\psi,$$

however, do not lend themselves to the formation of the desired complete system of eigenfunctions. They all have the form $\psi = e^{\frac{2\pi i}{h}r'p_r'}u(\theta',\varphi')$, and none of them is quadratically integrable. The eigenvalue spectrum, if any, must be continuous. The eigendifferentials

$$\Delta_{p_r'}^{\eta}\psi = \int_{d_r'}^{p_r'+\eta} u e^{\frac{2\pi i}{h}r'p_r''}dp_{r''} = \frac{h}{2\pi i r'}\left[e^{\frac{2\pi i}{h}r'\eta} - 1\right]e^{\frac{2\pi i}{h}r'p_r'}u$$

are quadratically integrable but do not vanish at the origin and do not belong to a linear manifold with respect to which $\frac{h}{2\pi i}\frac{\partial}{\partial r'}$ is Hermitian. Hence the operator $\frac{h}{2\pi i}\frac{\partial}{\partial r'}$ does not define a type 1 operator. As it is not a multiplication operator we can reasonably infer that it does not define a true dynamical variable in the quantum-mechanical sense.

It is of interest to note, however, that this operator does share several of the important properties of true dynamical variables—perhaps we might call it a *quasi-variable*. For example, we can speak of the "mean value" of p_r for an assemblage of a state described by the probability amplitude $(r',\theta',\varphi'|A) = w_A$ if we define the mean by

$$\bar{p_r} = \int\int\int w_A^* \frac{h}{2\pi i}\frac{\partial}{\partial r'} w_A dr'd\theta'd\varphi'. \tag{39.9}$$

This mean value is real, and its time derivative is given by (38.6). To be specific, we note that

$$p_r^{\binom{x}{x}} = (T^{-1})^{\binom{x\,y\,z}{r\,\theta\,\varphi}}\frac{h}{2\pi i}\frac{\partial}{\partial r'}T^{\binom{r\,\theta\,\varphi}{x\,y\,z}} = \Theta^{\binom{x\,y\,z}{r\,\theta\,\varphi}}\left[\frac{1}{r'}\frac{h}{2\pi i}\frac{\partial}{\partial r'}r'\right]\Theta^{\binom{r\,\theta\,\varphi}{x\,y\,z}}.$$

Hence, if we use wave functions normalized in Cartesian coordinates, Eq. (38.6) becomes in this case

$$\frac{d\bar{p_r}}{dt} = \int\int\int_{\infty}\Psi_X^*\left\{H^{\binom{x}{x}}\frac{1}{r'}\frac{\partial}{\partial r'}r' - \frac{1}{r'}\frac{\partial}{\partial r'}r'H^{\binom{x}{x}}\right\}\Psi_X dx'dy'dz'$$

$$= 2\frac{\overline{\mathcal{E}^2}}{r^3} - \frac{\overline{\partial V}}{\partial r}. \tag{39.10}$$

Although the definition (39.7) does not always yield a conjugate which is a true dynamical variable, it is the accepted definition for all cases in

which the coordinate q_k has a purely continuous spectrum.   The reader will note that, since q_k does not uniquely determine the q coordinate system to which .(39·7) directly refers, it is in general possible to use that equation to define several different quantities p_k, each of which is the momentum conjugate to q_k in an appropriate coordinate system.

Let us now assume that the real dynamical variable q_1, which in conjunction with $q_2, \cdots, q_\lambda$ forms a normal complete set of commuting variables, has a conjugate momentum p_1 defined by (39·7) which is a true dynamical variable.   The eigenfunctions of p_1 in q' space must be of the form

$$(q_1', \cdots, q_\lambda'|p_1', \cdots) = e^{\frac{2\pi i}{h}p_1'q_1'}u(q_2', \cdots, q_\lambda'). \quad (39\cdot11)$$

Simultaneous eigenfunctions of $p_1, q_2, \cdots, q_\lambda$ have the form

$$(q_1', \cdots, q_\lambda'|p_1', q_2'', \cdots, q_\lambda'')$$
$$= Ae^{\frac{2\pi i}{h}p_1'q_1'}\delta(q_2' - q_2'')\delta(q_3' - q_3'') \cdots \delta(q_\lambda' - q_\lambda''), \quad (39\cdot12)$$

A denoting a normalization factor.   They define a transformation operator $T^{\left(\alpha\atop q\right)}$ which carries probability amplitudes in q' space over into corresponding probability amplitudes in a space in which the coordinates are

$$\alpha_1 = p_1, \quad \alpha_2 = q_2, \quad \alpha_3 = q_3, \cdots, \quad \alpha_\lambda = q_\lambda.$$

The operator $T^{\left(\alpha\atop q\right)}$ has the specific form

$$T^{\left(\alpha\atop q\right)}(q'|) = \overline{\sum}_{q'}(\alpha'|q')(q'|)$$
$$= \overline{\sum}_{q'}Ae^{-\frac{2\pi i}{h}p_1'q_1'}\delta(q_2' - q_2'') \cdots \delta(q_\lambda' - q_\lambda'')(q'|). \quad (39\cdot13)$$

With the aid of $T^{\left(\alpha\atop q\right)}$ we can replace (39·7) by

$$p_1^{\left(q\atop q\right)} = (T^{-1})^{\left(q\atop q\right)}p_1'T^{\left(\alpha\atop q\right)}. \quad (39\cdot14)$$

Equations (39·11), (39·13), and (39·14) afford a generalization of the definition of the phrase "momentum quantum-mechanically conjugate to the coordinate q_k" to real dynamical variables q_k which are not positional coordinates and which do not have purely continuous spectra. An immediate consequence of this generalized definition is the rule that if β is the momentum quantum-mechanically conjugate to α, $-\alpha$ is the momentum quantum-mechanically conjugate to β. Thus the angular momentum $\mathcal{L}_z$ is the momentum quantum-mechanically conjugate to the angle φ in the coordinate system of Sec. 34d, and conversely $-\varphi$ is the "momentum" quantum-mechanically conjugate to the

"coordinate" $\mathcal{L}_z$ in the coordinate system $\mathcal{L}_i$, r_k, θ_k, α_k.   This rule forms an exact parallel of the corresponding rule for classical mechanics.

Unfortunately, however, it has not yet proved possible to derive the commutation rule (37·8) except on the basis of the definition (39·7). As this commutation rule is the usual basis of the applications of the idea of canonically conjugate variables in quantum mechanics, we can, for practical purposes, limit the conception to cases where the coordinate q has a purely continuous spectrum and its conjugate momentum p is defined by (39·7).

A still more drastic restriction of the class of canonical variables in quantum mechanics has been suggested by Dirac,[1] who purports to show that the relation $[p,q] = 1$ implies that p and q have continuous spectra extending from $-\infty$ to $+\infty$. The appearance of this statement[2] in the literature is rather mystifying in view of its immediate contradiction by the familiar and elementary example of Eq. (39·6). The angular momentum $\mathcal{L}_z$ has a purely discrete spectrum and the coordinate φ which defines the absolute azimuth of a system of particles with respect to the z axis has a continuous spectrum ranging from zero to 2π (according to the usual convention regarding the principal values), and yet $\left[\varphi, \dfrac{h}{2\pi i} \dfrac{\partial}{\partial \varphi} \right] = 1$.

An examination of Dirac's reasoning affords an illustration of the dangers of attempting to deal with quantum mechanics on a too formal and algebraic basis. As a corollary on (39·7) (dropping the unessential subscript k) we have the relation

$$p^{\left(q\atop q\right)}f(q') - f(q')p^{\left(q\atop q\right)} = \frac{df(q')}{dq'}. \tag{39·15}$$

In particular, it follows that for any number c,

$$p^{\left(q\atop q\right)}e^{icq'} = e^{icq'}p^{\left(q\atop q\right)} + \frac{hc}{2\pi}e^{icq'}. \tag{39·16}$$

Let us now apply this relation to an eigenfunction of p in q' space, which we designate as $\psi_{p'}$:

$$p^{\left(q\atop q\right)}(e^{icq'}\psi_{p'}) = \left(p' + \frac{ch}{2\pi}\right)(e^{icq'}\psi_{p'}). \tag{39·17}$$

From this equation Dirac draws the conclusion that we have only to multiply $\psi_{p'}$ by $e^{icq'}$, where c is now restricted to real values, in order to obtain an eigenfunction of p with the eigenvalue $p' + \dfrac{ch}{2\pi}$.   In other words, every real number is an eigenvalue of p'.   Complex eigenvalues are excluded on the ground that, if c contains an imaginary part, $e^{icq'}$ is not a "physically admissible" operator.

The argument breaks down, however, for real as well as imaginary values of c because it takes no account of the requirement that the eigenfunctions of p must belong to the Hermitian manifold of p, i.e., to D_p.   Identifying p with the angular momentum $\mathcal{L}_z$ and q with the angle φ, the domain D_p is defined by the requirement

[1] P. A. M. DIRAC, *P.Q.M.*, 1st ed., p. 54; 2d. ed., p. 94.

[2] Dirac's conclusion is apparently correct if the commutation rule $[p,q] = 1$ is assumed to be a valid matrix relation (*cf.* p. 367), as well as an operator relation.

that $\psi(\varphi + 2\pi) = \psi(\varphi)$. As we proved in Sec. 34 this boundary condition gives eigenfunctions of the form $\psi_m = e^{im\varphi}\chi(r_k,\theta_k,\alpha_k)$, where m is an integer and $\mathcal{L}_{z'} = mh/2\pi$. Evidently $e^{ic\varphi}\psi_m$ does not belong to D_p unless c is an integer.

39b. Functions of Non-commuting Linear Operators.

—The operator algebra of Sec. 37 permits us to define simple functions of non-commuting sets of linear operators, or dynamical variables, as well as of normal commuting sets. Thus the equation $\mathcal{L}^2 = \mathcal{L}_x^2 + \mathcal{L}_y^2 + \mathcal{L}_z^2$ defines $\mathcal{L}^2$ as a function of the elementary operators $\mathcal{L}_x$, $\mathcal{L}_y$, $\mathcal{L}_z$. Similarly,

$$\mathcal{L}_z = xp_y - yp_x$$

defines $\mathcal{L}_z$ as a function of the operators x, y, p_x, p_y. Such a function $F(\alpha,\beta, \cdots)$ of non-commuting operators has in general very different properties, however, from the previously defined functions of sets of normal commuting dynamical variables. As the "argument operators" α, β, $\cdots$ have no simultaneous eigenfunctions—at least not a complete set—there can be no general simple relation between the eigenfunctions of F and those of its arguments.

The concept of differentiation is applicable to functions of operators and can be included among the analytical processes available for building up such functions. Let I denote the identical operator which carries any function over into itself. The partial derivative of $F(\alpha,\beta, \cdots)$ with respect to α is then defined by the equation

$$\frac{\partial F}{\partial \alpha} = \lim_{a \to 0} \left[\frac{F(\alpha + aI, \cdots) - F(\alpha, \cdots)}{a} \right]. \quad (39\cdot18)$$

It is readily proved that on the basis of this definition all the ordinary rules of differentiation apply to the differentiation of operator functions except that one must have due regard to the order of the factors in dealing with products. Thus

$$\frac{\partial(F + G)}{\partial \alpha} = \frac{\partial F}{\partial \alpha} + \frac{\partial G}{\partial \alpha}, \quad (39\cdot19)$$

$$\frac{\partial(FG)}{\partial \alpha} = \frac{\partial F}{\partial \alpha}G + F\frac{\partial G}{\partial \alpha}, \quad (39\cdot20)$$

$$\frac{d}{d\alpha}(\alpha^n) = n\alpha^{n-1}. \quad (39\cdot21)$$

If the reciprocal of an operator α exists, Eq. (39·20) shows that

$$\frac{\partial}{\partial \alpha}(\alpha^n\alpha^{-n}) = \frac{\partial\alpha^n}{\partial \alpha}\alpha^{-n} + \alpha^n\frac{\partial\alpha^{-n}}{\partial \alpha} = 0.$$

It follows at once that (39·21) holds for negative values of n as well as for positive values.

The process of differentiation takes on a particularly simple form when we have to do with a function of one or more pairs of canonically con-

jugate operators. It makes no difference whether these operators are true dynamical variables or quasi-variables like the radial momentum p_r. If $F(p,q)$ denotes a function of the canonically conjugate pairs of operators $q_1,p_1,\ q_2p_2,\ \cdot\ \cdot\ \cdot\ ,q_f,p_f$ built up from its arguments by addition, multiplication, and division, the partial derivatives of F with respect to these arguments are

$$\frac{\partial F}{\partial q_k} = \frac{2\pi i}{h}(p_k F - F p_k) = [F,p_k], \tag{39.22}$$

$$\frac{\partial F}{\partial p_k} = \frac{2\pi i}{h}(F q_k - q_k F) = [q_k,F]. \tag{39.23}$$

To prove this statement we note first that the commutation rule

$$[q_l,p_k] = -[p_k,q_l] = \delta_{kl}$$

implies that (39.22) and (39.23) hold when we identify F with any one of the $2f$ basic operators $q_1,\ \cdot\ \cdot\ \cdot\ ,p_f$. But if the theorem be true for any two functions F and G it holds also for their sum and product. Thus,

$$\frac{\partial}{\partial q_k}(FG) = \frac{\partial F}{\partial q_k}G + F\frac{\partial G}{\partial q_k} = \frac{2\pi i}{h}[(p_k F - F p_k)G + F(p_k G - G p_k)]$$

$$= [FG,p_k].$$

The rules hold true even for negative powers of the q's if they are multiplicative operators representing positional coordinates. Thus by induction the rules (39.22) and (39.23) hold for any function F which can be built up legitimately by the operations of addition, multiplication, and division.

39c. An Operator Form of Hamilton's Equations of Motion.—
Equation (38.6) suggests the possibility of defining an operator $d\alpha/dt$ when α does not contain the time by

$$\frac{d\alpha}{dt} \equiv [\alpha,H], \tag{39.24}$$

thus insuring that when Ψ is a solution of the second Schrödinger equation

$$\frac{\overline{d\alpha}}{dt} = \int_\infty \Psi^*[\alpha,H]\Psi d\tau = \frac{d\bar{\alpha}}{dt}. \tag{39.25}$$

Let us assume that the operator α is Hermitian with respect to the linear manifold D of physically admissible wave functions and that the transform of any function in D by α belongs to the Hermitian manifold of H. Then, from the relation

$$([\alpha,H]\psi_1,\psi_2) = \frac{2\pi i}{h}\{(\alpha\psi_1,H\psi_2) - (H\psi_1,\alpha\psi_2)\} = (\psi_1,[\alpha,H]\psi_2),$$

it follows that $[\alpha,H]$ is Hermitian with respect to D.

If H is now assumed to be given as a function of a set of generalized positional coordinates and their conjugate momenta, and if we make use of the partial derivative formulas (39·22) and (39·23), we obtain, as a formal parallel to the classical equations of Hamilton,

$$\frac{dp_k}{dt} = -\frac{\partial H(p,q)}{\partial q_k}, \qquad \frac{dq_k}{dt} = \frac{\partial H(p,q)}{\partial p_k}. \qquad k = 1, 2, \cdots \quad (39\cdot26)$$

These formulas are largely definition, but, when used in connection with Eq. (39·25), form the basis for a quantum-mechanical derivation of the basic Hamiltonian equations of classical mechanics.

To this end we apply the equations of motion to a sharply defined wave packet, *i.e.*, to a wave function representing a pure case assemblage[1] of heavy systems so prepared that the uncertainty of each of the coordinates and momenta is very small compared with the absolute value of the quantity in question. If α denotes a coordinate with a continuous range of eigenvalues

$$\alpha\Psi = \alpha\int c(\alpha')\Psi_{\alpha'}d\alpha' = \int c(\alpha')\alpha'\Psi_{\alpha'}d\alpha'.$$

Since the packet is sharply defined, $c(\alpha')$ is appreciably different from zero only in a very narrow interval enclosing the average value $\bar{\alpha}$. Hence we have to a close approximation

$$\alpha\Psi = \bar{\alpha}'\int c(\alpha')\Psi_{\alpha'}d\alpha' = \bar{\alpha}\Psi. \qquad (39\cdot27)$$

Thus Ψ becomes an approximate eigenfunction for all the p's and q's, and for any function of them. All the coordinates and momenta commute to this approximation and the average value of any function of the coordinates and momenta becomes equal to the same function of the average values of its arguments. In particular $\overline{H(p,q)}$ goes over into $H(\bar{p},\bar{q})$ and $\overline{\partial H(p,q)/\partial q_k}$ goes over into $\partial H(\bar{p},\bar{q})/\partial\bar{q_k}$. Equations (39·25) and (39·26) now yield

$$\left.\begin{aligned} \frac{d\overline{p_k}}{dt} &= -\int_\infty \Psi^*\frac{\partial H(p,q)}{\partial q_k}\Psi d\tau = -\frac{\partial H(\bar{p},\bar{q})}{\partial\overline{q_k}}, \\ \frac{d\overline{q_k}}{dt} &= \frac{\partial H(\bar{p},\bar{q})}{\partial p_k}. \end{aligned}\right\} \qquad (39\cdot28)$$

Identifying the classical values of the p's and q's with the mean values for the packet, we obtain the classical canonical equations of Hamilton and thereby justify our quantum-mechanical definition of the momentum operator conjugate to the generalized positional coordinate q_k. Of course, we have not proved that $\frac{h}{2\pi i}\frac{\partial}{\partial q_k}$ is the *only* Hermitian operator whose mean value in the limiting case of a sharply defined wave packet

[1] Our results apply also to suitably defined mixtures.

will go over into the classical momentum conjugate to q_k. Whether other operators of this type exist or not is of no great importance.

40. SYMMETRY PROPERTIES OF THE WAVE EQUATION

40a. Symmetry Properties in General.—The conservation of angular momentum is one of several important symmetry properties of a free atomic system or of its Hamiltonian operator. These symmetry properties are of the greatest practical importance as they are responsible for most of the degeneracy of the energy levels, for the classification of the levels into non-combining term systems, etc.

In order to define the symmetry properties we assume a system of coordinates $x_1, x_2, \cdots, x_f$ and consider the substitution or transformation

$$\left. \begin{aligned} x_1 &\rightarrow \tilde{x}_1 = \varphi_1(x_1, \cdots, x_f), \\ x_2 &\rightarrow \tilde{x}_2 = \varphi_2(x_1, \cdots, x_f), \\ &\cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot, \\ x_f &\rightarrow \tilde{x}_f = \varphi_f(x_1, \cdots, x_f). \end{aligned} \right\} \tag{40.1}$$

Let R denote the operation of replacing x_1 by $\tilde{x}_1$ or $\varphi_1(x)$, x_2 by $\tilde{x}_2$ or $\varphi_2(x)$, $\cdots$. Thus

$$Rf(x_1, \cdots, x_f) = f[\varphi_1(x), \varphi_2(x), \cdots, \varphi_f(x)] \equiv \tilde{f}(x). \tag{40.2}$$

Here $\tilde{f}(x)$ is simply the new function of x obtained by letting R act on $f(x)$. Let us now define the operator $\tilde{H}$ by the equation

$$RH\psi(x) = \tilde{H}\tilde{\psi}(x) = \tilde{H}R\psi. \tag{40.3}$$

We may say that $\tilde{H}$ is the operator into which H is carried by the substitution R. If $\tilde{H}$ is the same as the original operator H we say that H is *invariant* under the substitution R. The property of being invariant with respect to the substitution R is a symmetry property of H. It means that H commutes with R. If R has an adjoint manifold which includes D and a complete set of eigenfunctions it is a proper dynamical variable which is conserved during the natural motion of the system. In other words R is an integral of the motion described by the Schrödinger wave equation.

As a simple example we consider the one-dimensional case where

$$H = -\frac{h^2}{8\pi^2\mu} \frac{d^2}{dx^2} + V(x).$$

Then, if R is any substitution (40.2),

$$RH\psi = -\frac{h^2}{8\pi^2\mu} \frac{d^2}{d\varphi^2}\psi[\varphi(x)] + V[\varphi(x)]\psi[\varphi(x)],$$

or

$$\tilde{H} = -\frac{h^2}{8\pi^2\mu}\left[\left(\frac{dx}{d\varphi}\right)^2 \frac{d^2}{dx^2} + \frac{d^2x}{d\varphi^2} \frac{d}{dx}\right] + \tilde{V}(x)$$

In order that $\tilde{H}$ shall be identical with H it is necessary and sufficient that φ has the special form

$$\varphi(x) = a \pm x,$$

where a is a constant, and that

$$\tilde{V} = V(a \pm x) = V(x).$$

If V is not constant, it must either be periodic with period a or it must be symmetrical about the point $x = \dfrac{a}{2}$.

Let us consider the latter case, of which the Planck linear oscillator is a special example. Taking the point of symmetry as the origin, the transformation R becomes the operation of replacing x by $-x$, or reversing the direction of the x axis. We designate this particular transformation or operator by the symbol U. U is Hermitian, since

$$\int_{-\infty}^{+\infty} [u^*(x)Uv(x) - v(x)Uu^*(x)]dx$$
$$= \int_{-\infty}^{+\infty} [u^*(x)v(-x) - v(x)u^*(-x)]dx = 0. \quad (40 \cdot 4)$$

The integral vanishes because the integrand is an odd function of x, *i.e.*, a function which has opposite values at x and $-x$. Since $U^2 = 1$, and $U = U^\dagger$, U belongs to the class of unitary operators defined in Sec. 36c.

An operator such as U which is both Hermitian and unitary can have but two eigenvalues ± 1, for, since $U^2 = 1$, each eigenvalue must be a real square root of unity. We shall use the term *symmetry values* for the eigenvalues of these operators. Eigenfunctions of any of them are said to be "symmetric" or "antisymmetric" with respect to the operator in question according as the symmetry value is $+1$ or -1.

An *even* function of x is by definition a solution of the equation $U\psi = \psi$ and, if quadratically integrable, can be regarded as an eigenfunction of U with the eigenvalue $+1$. Similarly a quadratically integrable *odd* function of x is an eigenfunction of U with the eigenvalue -1. An arbitrary quadratically integrable function $f(x)$ can always be resolved into the sum of two such eigenfunctions. To do so one has only to form the functions

$$f^{(+)} = \tfrac{1}{2}(f + Uf), \qquad f^{(-)} = \tfrac{1}{2}(f - Uf). \quad (40 \cdot 5)$$

Then

$$Uf^{(+)} = f^{(+)}, \qquad Uf^{(-)} = -f^{(-)}, \quad (40 \cdot 6)$$
$$f = f^{(+)} + f^{(-)}. \quad (40 \cdot 7)$$

Clearly U is a real dynamical variable.

Not only do U and H commute. They also satisfy the conditions **(b)** and **(c)** of Sec. 37c (p. 285) needed to establish the fact that they form a normal commuting pair. Thus U has a purely discrete spectrum and it is possible to approximate any of its eigenfunctions with arbitrary precision by means of another eigenfunction of the same eigenvalue which belongs to the Hermitian domain of H. Hence every eigenfunction of H must be expansible in terms of simultaneous eigenfunctions of U and H. Since the one-dimensional energy equation is nondegenerate, this means that every eigenfunction of H is also an eigenfunction of U. In other words, the eigenfunctions of H are either even or odd functions of x. The operator U is thus a function of H in the sense of the definition on p. 281.

40b. The Reflection Operators.—In the theory of atomic and molecular structure we meet with a number of operators similar in their properties to U. Consider the substitution

$$x_i \rightarrow \tilde{x}_i = -x_i, \qquad y_i \rightarrow \tilde{y}_i = y_i, \qquad z_i \rightarrow \tilde{z}_i = z_i, \qquad (40\cdot8)$$

where i ranges from 1 to f in a problem involving f particles. This transformation carries each particle over into a position previously occupied by its mirror image in the y,z plane. Hence the transformation can be described as a reflection and its operator R_x as a *reflection operator*. R_x is defined by the above transformation equations, or by

$$R_x f(x_1,y_1,z_1,x_2,y_2, \cdots ,z_f) = f(-x_1,y_1,z_1,-x_2,y_2, \cdots ,z_f). \quad (40\cdot9)$$

The corresponding operators denoting reversal of the signs of the y and z coordinates will be designated by R_y and R_z, respectively. Another operator which shares most of their properties is the product

$$K = R_x R_y R_z,$$

which can also be defined by the equation

$$Kf(x_1,y_1,z_1, \cdots ,z_f) = f(-x_1,-y_1,-z_1, \cdots ,-z_f). \quad (40\cdot10)$$

Let R denote an arbitrary member of the set of four operators R_x, R_y, R_z, K. It is easy to see by the argument used for U that R is Hermitian and unitary. Since an arbitrary function can always be resolved into the sum of symmetric and antisymmetric functions by the device indicated in Eqs. (40·5), (40·6), and (40·7), these operators represent true dynamical variables.

R_x, R_y, R_z, K, and $\mathcal{L}^2$ all commute with the fundamental Hamiltonian of Eq. (32·1) and with each other. They all have purely discrete spectra and in fact are readily proved to unite with H to form a normal commuting set of dynamical variables. Hence it is possible to expand an arbitrary wave function in terms of simultaneous eigenfunctions of H, $\mathcal{L}^2$,

R_x, R_y, R_z, K. In the case of the two-particle problem of Sec. 28, for example, we accomplish this by writing

$$\psi = \sum_{n=1}^{\infty} \sum_{l=0}^{n-1} \sum_{m=0}^{l} R_{nl}(r)\Theta_{lm}(\theta)(a_{nlm}\cos m\varphi + b_{nlm}\sin m\varphi). \quad (40\cdot11)$$

$\mathcal{L}_z$ commutes with R_z and K but not with R_x and R_y. Hence we cannot expand in terms of simultaneous eigenfunctions of $\mathcal{L}_z$ and R_x or R_y but can expand in terms of simultaneous eigenfunctions of $\mathcal{L}_z$, R_z, and K.

40c. The Rotation Operator.—Consider next a many-particle problem in which the potential energy V is independent of a rotation of the system as a rigid whole about the z axis. Such a rotation is described by the transformation

$$F(\omega):\begin{cases} x_i \rightarrow \tilde{x}_i = x_i \cos \omega - y_i \sin \omega, \\ y_i \rightarrow \tilde{y}_i = x_i \sin \omega + y_i \cos \omega, \\ z_i \rightarrow \tilde{z}_i = z_i. \end{cases} \quad (40\cdot12)$$

V is by hypothesis invariant with respect to F, and the symmetry of the Laplacian operator insures that the complete energy operator H shall commute with F.

The operator F has a reciprocal F^{-1} obtained by reversing the sign of the angle ω. Clearly the application of F or F^{-1} to the complete integrand of an integral extended over all coordinate space cannot affect the value of the integral. If ψ and χ are any two quadratically integrable functions, it follows that

$$(\psi, F^{-1}\chi) = \int_{\infty} F[\psi(F^{-1}\chi)^*]d\tau = (F\psi, \chi).$$

Thus F is unitary with respect to coordinate space and the class of all quadratically integrable functions.

Expanding the function $F\psi = \tilde{\psi}$ in power series in ω, we obtain[1]

$$F\psi = \psi(x) + \omega\left(\frac{\partial\tilde{\psi}}{\partial\omega}\right)_{\omega=0} + \frac{\omega^2}{2!}\left(\frac{\partial^2\tilde{\psi}}{\partial\omega^2}\right)_{\omega=0} + \cdots. \quad (40\cdot13)$$

But

$$\left(\frac{\partial\tilde{\psi}}{\partial\omega}\right)_{\omega=0} = \sum_j\left[\frac{\partial\tilde{\psi}}{\partial\tilde{x}_j}\frac{\partial\tilde{x}_j}{\partial\omega} + \frac{\partial\tilde{\psi}}{\partial\tilde{y}_j}\frac{\partial\tilde{y}_j}{\partial\omega} + \frac{\partial\tilde{\psi}}{\partial\tilde{z}_j}\frac{\partial\tilde{z}_j}{\partial\omega}\right]_{\omega=0}$$

$$= \sum_j\left\{\left[\tilde{x}_i\frac{\partial}{\partial\tilde{y}_i} - \tilde{y}_i\frac{\partial}{\partial\tilde{x}_i}\right]\tilde{\psi}\right\}_{\omega=0}.$$

and

$$\left(\frac{\partial^n\tilde{\psi}}{\partial\omega^n}\right)_{\omega=0} = \left(\frac{2\pi i\mathcal{L}_z}{h}\right)^n\psi.$$

[1] The expansion is legitimate if $\psi(x)$ is of class D, for in that case $\psi(x)$ is analytic in the Cartesian coordinates except at points where V is infinite. Hence it is analytic in φ and it follows that $F(\omega)\psi$ is analytic in ω.

Hence (40·13) is equivalent to

$$F(\omega) = \sum_{k=0}^{\infty} \frac{(2\pi i \omega \mathfrak{L}_z / h)^k}{k!} = e^{\frac{2\pi i}{h} \omega \mathfrak{L}_z}. \tag{40·14}$$

Equation (40·13) leads us to a functional relationship between the rotation operator $F(\omega)$ and the familiar angular momentum operator $\mathfrak{L}_z$. Since $\mathfrak{L}_z$ is a real dynamical variable, it follows that $F(\omega)$ is a dynamical variable. If $\mathfrak{L}_z$ had not been defined previously, we could define it by

$$\mathfrak{L}_z \psi = \frac{h}{2\pi i} \left[\frac{\partial (F\psi)}{\partial \omega} \right]_{\omega = 0}. \tag{40·15}$$

Moreover, the Hermitian character of $\mathfrak{L}_z$ and the fact that it is an integral of the motion are simple corollaries of the properties of $F(\omega)$. Thus we can say that the existence' of the family of rotation operators $F(\omega)$ which commute with H *generates* the integral of the motion $\mathfrak{L}_z$.

Let us next consider the case where V is invariant with respect to all rigid rotations of the axes. Then an operator representing any such rotation will commute with H. The most general rigid displacement of the axes can be accomplished by a single rotation through a suitably chosen angle ω about a suitably chosen axis. If the direction cosines of the axis are λ, μ, ν, the rotation is represented by the vector $\overrightarrow{\omega}$ whose components are $\omega_x = \lambda\omega$, $\omega_y = \mu\omega$, $\omega_z = \nu\omega$. To apply such a rotation to ψ, we can imagine a preliminary shift in reference axes from the primary x, y, z system to a system x', y', z' so chosen that z' has the direction of $\overrightarrow{\omega}$. If we now apply the operator $e^{\frac{2\pi i \omega}{h} \mathfrak{L}_{z'}}$ and then transform back to the x, y, z system we shall have the equivalent of a direct rotation of the x, y, z axes of the desired kind. But the operators $\mathfrak{L}_x$, $\mathfrak{L}_y$, $\mathfrak{L}_z$ are readily seen to transform like the components of a vector. Hence

$$\mathfrak{L}_{z'} = \lambda\mathfrak{L}_x + \mu\mathfrak{L}_y + \nu\mathfrak{L}_z.$$

Consequently the desired substitution is effected directly on the original wave function in x, y, z coordinates by the dynamical-variable operator

$$G(\overrightarrow{\omega}) = e^{\frac{2\pi i}{h}\omega(\lambda\mathfrak{L}_x + \mu\mathfrak{L}_y + \nu\mathfrak{L}_z)} = e^{\frac{2\pi i}{h}[\omega_x\mathfrak{L}_x + \omega_y\mathfrak{L}_y + \omega_z\mathfrak{L}_z]}.$$

The operator $G(\overrightarrow{\omega})$ is not to be identified with the product of the three non-commutative operators $e^{\frac{2\pi i}{h}\omega_x\mathfrak{L}_x}$, $e^{\frac{2\pi i}{h}\omega_y\mathfrak{L}_y}$, and $e^{\frac{2\pi i}{h}\omega_z\mathfrak{L}_z}$, but with the power-series expansion $\sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{2\pi i}{h}\right)^k (\omega_x\mathfrak{L}_x + \omega_y\mathfrak{L}_y + \omega_z\mathfrak{L}_z)^k$. $G(\overrightarrow{\omega})\psi$ is an

analytic function of ω_x, ω_y, ω_z. If we expand $\overrightarrow{G(\omega)}$ as a formal multiple power series in ω_x, ω_y, ω_z, the coefficient of each term in the expansion is a homogeneous form in $\mathcal{L}_x$, $\mathcal{L}_y$, $\mathcal{L}_z$ which is either Hermitian, or becomes so when multiplied by i. Thus $\overrightarrow{G(\omega)}$ generates a large multiplicity of Hermitian operators (including $\mathcal{L}_x$, $\mathcal{L}_y$, $\mathcal{L}_z$) which commute with H because $\overrightarrow{G(\omega)}$ commutes with H.

The operators $\overrightarrow{G(\omega)}$ form a continuous "group" in the technical sense of the word.[1] They are a subgroup of the more general rotation-reflection group which includes also all operators of the type R_x, R_y, R_z. The operator $F(\pi) = G(0,0,\pi)$ reverses the direction of the x and y axes and is therefore equal to the product of R_xR_y. It follows that the product of $F(\pi)$ into K is R_z. Since $F(\pi)$ is a function of $\mathcal{L}_z$, R_z is a function of the commuting operators $\mathcal{L}_z$ and K. Similarly R_x and R_y are functions of $\mathcal{L}_x$ and K and of $\mathcal{L}_y$ and K, respectively. *Consequently every substitution in the rotation-reflection group can be effected with the aid of K and $\overrightarrow{G(\omega)}$.*

Since the application of any member of this group, say $\overrightarrow{G(\omega)}$, to the integrand of a definite integral extended over all configuration space cannot change the value of the integral, we have

$$(H\psi,\chi) = \int_\infty G(\chi^*H\psi)d\tau = \int_\infty (HG\psi)(G\chi)^*d\tau = (HG\psi,G\chi).$$

Consequently *every such operator transforms the Hermitian manifold of H into itself.*

If K and every $\overrightarrow{G(\omega)}$ commute with an operator α, it follows that every dynamical variable generated by the group of substitutions commutes with α. This is true of the dynamical variables K and $\mathcal{L}^2$.

40d. The Permutation Operators.—In addition to the invariance of the Hamiltonian operator of a free atomic system with respect to substitutions of the rotation-reflection group there is an invariance with respect to a second set of substitutions which permutes the coordinates of equivalent particles, *i.e.*, permutes the coordinates of different electrons of the system, or of different protons, as the case may be. Symmetry operators of this second type also generate dynamical variables which are integrals of the motion.

The simplest class of permutation operators are the *interchange operators or transpositions* which permute the coordinates of two particles only. Thus P_{ij} is the operator which performs the transformation

$$x_i \rightarrow x_j, \qquad y_i \rightarrow y_j, \qquad z_i \rightarrow z_j,$$
$$x_j \rightarrow x_i, \qquad y_j \rightarrow y_i, \qquad z_j \rightarrow z_i,$$

---

[1] *Cf., e.g.*, Carl Eckart, *Rev. Mod. Phys.* **2**, 305 (1930) or the treatises on quantum mechanics from the group-theory point of view by Wigner, Weyl, and van der Waerden.

and so interchanges the positions of the ith and jth particles. It is convenient in discussing the permutations to designate the complete set of Cartesian coordinates of the ith particle by ξ_i. With this notation we can say that any permutation P effects a rearrangement of the arguments $\xi_1, \cdots, \xi_n$ of the ψ function to which it is applied. Including the identical permutation, the number of different permutations which can be applied to a wave function describing a system composed of n equivalent particles is $n!$

The result of two successive permutations is a permutation which we call the product of the other two. It is immediately evident that every one of the more complicated permutations can be built up as the product of a number of simple interchange permutations. The associative law holds for permutation products and each permutation has a reciprocal, or inverse. Hence the permutations form a group like the rotation and reflection operators.

The application of a permutation operator to the integrand of a definite integral is like the application of a rotation operator to such an integrand in leaving the value of the integral unchanged. It follows (*cf.* p. 308) that the permutations transform the Hermitian manifold of H into itself. The reader will also verify that they transform class D into itself.

Let P^{-1} denote the reciprocal of P. By the above rule

$$\int_\infty \psi_1{}^* P\psi_2 d\tau = \int_\infty (P^{-1}\psi_1{}^*)(P^{-1}P\psi_2)d\tau = \int_\infty \psi_2 P^{-1}\psi_1{}^* d\tau.$$

Thus P^{-1} is adjoint to P, and P has a unitary manifold which includes all quadratically integrable functions.

The interchange operators P_{ij} have the additional property that $P_{ij}{}^2 = 1$. Hence they are Hermitian as well as unitary and each has a complete system of eigenfunctions with the eigenvalues ± 1. The Hermitian manifold of any interchange operator includes class D so that these operators are real dynamical variables of type 1. Although the general permutations are not Hermitian, they also can be proved to have complete sets of eigenfunctions with discrete eigenvalues and are proper dynamical variables. Every permutation when raised to a suitable minimum power n, less than, or equal to, the number of permuted particles, gives unity.[1] Hence every eigenvalue is an nth root of unity, and is in general complex. Thus the eigenvalues are all of unit absolute value in accordance with the general theorem of Sec. 36j, p. 275.

In order to apply a permutation P to $F\psi(\xi_1,\cdots,\xi_n)$, where F denotes a function of the operators for the coordinates and momenta, we have to permute the subscripts in the expression for F as well as

[1] This is a general property of groups containing a finite number of elements (*cf.*, *e.g.*, E. Wigner, *Gruppentheorie und Ihre Anwendung auf die Quantenmechanik der Atomspektren*, p. 65, Braunschweig, 1931.

in the expression for ψ. If F is symmetrical in the coordinates and momenta of all equivalent particles, and if P permutes the subscripts of equivalent particles only, it follows that P commutes with F. In particular, P commutes with the Hamiltonian operator of any atomic system when the permuted particles are all of the same species. It follows that the permutation operators are integrals of the motion like the rotation-reflection operators.

The reciprocal of the product $\alpha\beta$ is $\beta^{-1}\alpha^{-1}$. If α and β are unitary and convert class D functions into class D functions, we have

$$(\alpha\beta\psi,\ \chi) = (\beta\psi,\ \alpha^{-1}\chi) = (\psi,\ \beta^{-1}\alpha^{-1}\chi) = (\psi,\ (\alpha\beta)^{-1}\chi),$$

provided that ψ and χ are physically admissible. Hence the product of any member of the permutation group into any member of the rotation-reflection group is a unitary operator. In fact it is easy to see that the operators of these two groups, together with their products, can be combined into a single unitary substitution group every element of which commutes with H in the case of a free atomic system. This group of unitary integrals is called the *group of the Schrödinger equation*, or the *symmetry group of configuration space*.[1]

The complete impossibility of distinguishing one electron from another implies the impossibility of measuring any dynamical variable which does not commute with all the permutation operators (*cf.* Sec. 42*b*). The space-symmetry integrals R_x, R_y, R_z, $\cdots$, $\mathcal{L}_z$ do commute with the permutation group and are measurable. On the other hand the P's themselves are not symmetrical and do not commute with each other in general. For example,

$$P_{12}P_{23}\psi(\xi_1,\xi_2,\xi_3) = \psi(\xi_2,\xi_3,\xi_1); \qquad P_{23}P_{12}\psi(\xi_1,\xi_2,\xi_3) = \psi(\xi_3,\xi_1,\xi_2).$$

40e. Degeneracy and the Integrals of the Schrödinger Equation.— We saw in Sec. 37*c* that if two type 1 dynamical variables α and β commute, the application of β to an eigenfunction of α, say φ_n, frequently gives a new eigenfunction of α, linearly independent of φ_n and belonging to the same eigenvalue. Even if β is not a type 1 dynamical variable, $\beta\varphi_n$ is an eigenfunction of the type specified provided that it belongs to the Hermitian manifold of α and is not a multiple of φ_n.

Every element U of the group of the Schrödinger equation commutes with H and transforms the Hermitian manifold of H into itself. If ψ_{lk} is a discrete eigenfunction of H with the eigenvalue E_k, it follows that

$$HU\psi_{lk} = UH\psi_{lk} = E_k U\psi_{lk},$$

---

[1] Strictly speaking the permutation group which commutes with a Hamiltonian depends on the number of particles of different species in the system under consideration. Also the complete rotation-reflection group fails to commute with H if an atom is subject to an external electric or magnetic field. Hence the group of the Schrödinger equation is variable according to the nature of H.

and that $U\psi_{lk}$ is also an eigenfunction of H with the eigenvalue E_k. $U\psi_{lk}$ will be linearly independent of ψ_{lk} unless the latter function is itself an eigenfunction of U. Unless ψ_{lk} is a simultaneous eigenfunction of *all* the operators in the group of the Schrödinger equation, some of the transforms will necessarily be linearly independent of ψ_{lk}, showing that E_k is degenerate.

If degeneracy is to result from the existence of an integral T of the Schrödinger equation, whether unitary or not, it is necessary that there shall also exist a second integral, say S, which does not commute with T. The condition is a corollary on the second theorem of p. 287. From that theorem it follows that if all integrals of the Schrödinger equation commute with any one of them, the operator in question must be a function of the energy operator. But a function of the energy cannot transform an eigenfunction of H into a linearly independent function. Hence every operator which can be said to produce degeneracy must commute with H and at the same time fail to commute with a second integral, S.

The existence of two non-commuting integrals T and S is not sufficient of itself to establish degeneracy, but becomes so if the integrals T and S have the property of transforming eigenfunctions of H into functions which belong to the Hermitian manifold of H. We state this conclusion in the form of a theorem.

Theorem: Let T and S denote a pair of integrals of the Schrödinger equation and let ψ_{lk} denote an eigenfunction of H with the eigenvalue E_k. If $(TS - ST)\psi_{lk} \neq 0$, and if $T\psi_{lk}$ and $S\psi_{lk}$ belong to the Hermitian manifold of H, it follows that the energy level E_k is degenerate and that there exist two or more simultaneous eigenfunctions of H and T, or H and S, which belong to E_k, but to different eigenvalues of T, or S, as the case may be.

Proof: It follows from the hypotheses that $T\psi_{lk}$ and $S\psi_{lk}$ are eigenfunctions of H with the eigenvalue E_k (*cf.* Sec. 37c). Since $(TS - ST)\psi_{lk} \neq 0$, ψ_{lk} cannot be a simultaneous eigenfunction of all three of the operators H, T, S. Hence one of the functions $T\psi_{lk}$ and $S\psi_{lk}$ must be linearly independent of ψ_{lk} and E_k must be degenerate. We assume that ψ_{lk} is not an eigenfunction of T—the argument is the same in form if we assume that ψ_{lk} is not an eigenfunction of S. Since E_k has at most a finite degeneracy (*cf.* Sec. 32k) we can express the eigenfunction $T\psi_{lk}$ as a linear combination of a finite orthonormal set of eigenfunctions, say ψ_{1k}, ψ_{2k}, $\cdots$, $\psi_{\lambda k}$. This set must contain at least two members. Making a similar expansion of the transform of each member of the set by T, we obtain λ equations similar to (37·11). By means of a linear transformation (*cf.* p. 284) we can always obtain a new orthonormal set of functions which are simultaneous eigenfunctions of H and T. These new functions cannot have a single common eigenvalue for T because ψ_{lk} is a linear combination of them and by hypothesis ψ_{lk} is not an eigenfunction of T. This completes the proof of our theorem.

Since the dynamical variables which make up the group of the Schrödinger equation for a free atomic system do not in general commute

with each other, and since they transform the Hermitian domain of H into itself (*cf.* Secs. 40*c* and 40*d*), it is clear that these operators will produce degeneracy. We shall refer to degeneracy of this kind as *symmetry degeneracy.*

In addition to symmetry degeneracy there are other but less important types. Thus in the case of hydrogenic atoms treated without relativity and spin corrections there is a degeneracy of the discrete energy levels with respect to $\mathcal{L}^2$, as we saw in Sec. 29. This degeneracy is peculiar to the two-particle problem with the Coulomb inverse-square law of attraction and will be called *Coulomb degeneracy.* It breaks down the $\mathcal{L}^2$ classification of the energy levels of the general two-particle problem and indicates the existence of an operator X, which commutes with H, but not with $\mathcal{L}^2$. Many such operators could be set up artificially in integral form, but no simple substitution has been proposed which accounts for this type of degeneracy as a symmetry property.[1]

Another type of systematic degeneracy not accounted for as an ordinary symmetry property is associated with the continuous part of the energy spectrum. Consider, for example, the case of a two-particle problem like that of the dumbbell model of the diatomic molecule (Sec. 28*h*). Whereas in the discrete portion of the energy spectrum every level is characterized by an individual eigenvalue of $\mathcal{L}^2$, the crowding together of the levels in the continuous spectrum makes every eigenvalue of $\mathcal{L}^2$ compatible with every eigenvalue of H. We call this *continuous-spectrum degeneracy.* It is not readily accounted for as a result of a symmetry property of the Hamiltonian.

In view of the circumstances giving rise to continuous-spectrum degeneracy we may say that, in the above case, $\mathcal{L}^2$ *is a function of H in the discrete portion of the energy spectrum but is independent of H in the continuous portion of the spectrum.* Within the discrete spectrum H and $\mathcal{L}_z$ suffice to form a complete set of normally commuting dynamical variables, but, when the continuous spectrum is taken into account, we must add $\mathcal{L}^2$ to H and $\mathcal{L}_z$ to get such a set. Exactly the same phenomena mark the passage from the discrete to the continuous spectrum in the case of a general many-electron atom.

Finally we have to consider cases of non-systematic degeneracy, or *accidental degeneracy,* due to the chance energy agreement of individual stationary states which do not belong together as a result of any general

---

[1] Recently, however, V. Fock, *Bull. de l'Académie des Sciences de L'URSS*, p. 179, 1935, has shown that the Schrödinger equation

$$H^{\left(p \atop p\right)}\phi(p') = E\phi(p')$$

for hydrogenic atoms in momentum space is identical with the integral equation for spherical harmonics in four-dimensional space. Thus the Coulomb degeneracy can be tied up with the symmetry properties of four-dimensional space.

rule. Examples of this kind are rare except in problems in which the Hamiltonian involves a continuously variable parameter, such as the strength of an electric or magnetic field. In such problems accidental degeneracy is apt to occur for certain specific values of the parameter in question at which a pair of energy levels are said to "cross each other." Since accidental degeneracy is exceptional, unpredictable, and easily recognized when it does occur, it can be ignored in a general discussion of normal systematic degeneracy.

It will now be seen that if we restrict our discussion to the discrete spectra of atomic systems with two or more electrons, we have to consider only the symmetry degeneracy associated with the group of the Schrödinger equation. This problem will be attacked on the basis of the *postulate that all eigenfunctions of H for any discrete energy level E_k are contained in the linear manifold formed from the transforms of any one eigenfunction by the operators in the unitary substitution group of the Schrödinger equation, together with their linear combinations.*[1] This is the fundamental postulate of all applications of group theory to atomic structure problems. It is plausible *a priori* and is substantiated both by experiment and by detailed calculation. It implies that the linear manifold of eigenfunctions belonging to E_k and obtained by application of the unitary group of the Schrödinger equation is independent of the choice of the initial wave function ψ_{lk}, so long as it belongs to E_k. This means that all eigenfunctions belonging to E_k have fundamentally the same symmetry with respect to the operations which define the group.

40f. The Normal Degeneracy of the Energy Levels of Free Atomic Systems.—Owing to the existence of symmetry degeneracy we know that even within the discrete spectrum the Hamiltonian H of such a system does not of itself constitute a complete set of normally commuting dynamical variables. Our first problem is then to choose additional dynamical variables which either belong to the group of the Schrödinger equation, or are related to it, and can be added to H in order to build up a complete normally commuting set. The degeneracy of any energy level E_k is then equal to the number of sets of eigenvalues of these additional operators which are compatible with E_k. Since all the permutations commute with all rotations and reflections, we can deal separately with the two subgroups of operators.

Consider first the rotation-reflection group. Finite rotations about different axes do not commute. This we know from geometry and from the fact that the corresponding components of angular momentum do not commute. On the other hand, all rotations about any one axis commute with each other and the corresponding component of angular momentum. Therefore, if we add to H any finite rotation, or any

---

[1] The phrase "linear combination" is to be interpreted to include such limiting cases as the derivative given in Eq. (40·15).

component of angular momentum, we shall have a normal commuting set of operators which does not commute with any independent member of the rotation group. This situation is unchanged if we pass from the pure rotation group to the rotation-reflection group, for K commutes with all the rotation and angular-momentum operators so that the commutation properties of $KG(\vec{\omega})$ are determined by those of G.

Let us arbitrarily choose $\mathcal{L}_z$ as the second member of our commuting set. If the system under consideration contains not more than two identical particles, there are only two permutations, one of which is the identical permutation and both commute with H and $\mathcal{L}_z$. In this case we infer that the set of operators $H,\mathcal{L}_z$ is complete. If there are more than two identical particles, it will be necessary to add one or more permutations in order to secure a complete set.

The operators $\mathcal{L}^2$ and K commute with every component of angular momentum and every operator of the rotation-reflection group. For example,

$$[\mathcal{L}^2,\mathcal{L}_x] = [\mathcal{L}_y^2,\mathcal{L}_x] + [\mathcal{L}_z^2,\mathcal{L}_x]$$
$$= \mathcal{L}_y\mathcal{L}_z + \mathcal{L}_z\mathcal{L}_y - \mathcal{L}_z\mathcal{L}_y - \mathcal{L}_y\mathcal{L}_z = 0.$$

$$[\mathcal{L}_x,K] = \left\{\sum_k\left[(-y_k)\left(-\frac{\partial}{\partial z_k}\right) - (-z_k)\left(-\left(\frac{\partial}{\partial y_k}\right)\right)\right]\right.$$

$$\left. - \sum_k\left[y_k\frac{\partial}{\partial z_k} - z_k\frac{\partial}{\partial y_k}\right]\right\}K = 0.$$

In fact, these operators commute with every member of the group of the Schrödinger equation. Since they are integrals of the motion, there exists a complete system of simultaneous eigenfunctions of $H,\mathcal{L}^2$, and F. Let ψ_{lk} denote an eigenfunction of this type associated with the discrete energy level E_k. From the postulate in italics on p. 313, it follows that every eigenfunction of H with the eigenvalue E_k can be derived from ψ_{lk} by taking linear combinations of transforms of ψ_{lk} by the various operators in the substitution group. But, since $\mathcal{L}^2$ and K commute with all members of the group of the Schrödinger equation, all eigenfunctions of H which belong to E_k are eigenfunctions of $\mathcal{L}^2$ and K with the same eigenvalues as ψ_{lk}. In other words, there is just one pair of eigenvalues of $\mathcal{L}^2$ and K which is compatible with any discrete eigenvalue of H. Thus $\mathcal{L}^2$ and K are functions of H insofar as concerns the discrete spectrum.

Thus the eigenvalues of $\mathcal{L}^2$ and K form convenient indices for the classification of energy levels. Stationary states which are symmetric with respect to K ($K' = +1$) are said to be *even*, while those which are antisymmetric are said to be *odd*. This classification of levels into even and odd types holds rigorously even when the electron spin is taken into account. The spin-orbit interactions partially spoil the $\mathcal{L}^2$ classifica-

tion, however, replacing it by a similar $\mathcal{g}^2$ classification, where $\overrightarrow{\mathcal{g}}$ denotes the resultant of the orbital and spin angular momenta. In the case of the two-particle problem without spin, where $\mathcal{L}^2$ has the form of Eq. (34·18), eigenfunctions of $\mathcal{L}^2$ with even values of l are symmetric with respect to K while those with odd values of l are antisymmetric. Thus K is a function of $\mathcal{L}^2$ for such systems.

Evidently the number of different $\mathcal{L}_z$ values compatible with a given energy level not degenerate with respect to $\mathcal{L}^2$ is not greater than the number of $\mathcal{L}_z$ values compatible with the corresponding value of L, viz., $2L + 1$ (cf. Sec. 34f, p. 234). Neither, on the other hand, can the degeneracy be less than $2L + 1$. To prove this we follow an argument used by Dirac.[1] Let $\psi_{E,L,M}$ be a simultaneous eigenfunction of $H,\mathcal{L}^2,\mathcal{L}_z$ with the discrete eigenvalues E, $L(L + 1)h^2/4\pi^2$, $Mh/2\pi$, respectively. Since $\mathcal{L}_x$ and $\mathcal{L}_y$ commute normally with H and $\mathcal{L}^2$, $(\mathcal{L}_x + i\mathcal{L}_y)\psi_{E,L,M}$ must either vanish identically or be an eigenfunction of H and $\mathcal{L}^2$ which, like $\psi_{E,L,M}$, belongs to E and L. The same remark applies to

$$(\mathcal{L}_x - i\mathcal{L}_y)\psi_{E,L,M}.$$

The commutation relations (38·15), (38·16), and (38·17) imply that

$$(\mathcal{L}_x + i\mathcal{L}_y)\mathcal{L}_z - \mathcal{L}_z(\mathcal{L}_x + i\mathcal{L}_y) = -\frac{h}{2\pi}(\mathcal{L}_x + i\mathcal{L}_y),$$

or that

$$\mathcal{L}_z(\mathcal{L}_x + i\mathcal{L}_y) = (\mathcal{L}_x + i\mathcal{L}_y)\left(\mathcal{L}_z + \frac{h}{2\pi}\right). \tag{40·16}$$

Hence

$$\mathcal{L}_z(\mathcal{L}_x + i\mathcal{L}_y)\psi_{E,L,M} = (M + 1)\frac{h}{2\pi}(\mathcal{L}_x + i\mathcal{L}_y)\psi_{E,L,M}. \tag{40·17}$$

We conclude that $(\mathcal{L}_x + i\mathcal{L}_y)\psi_{E,L,M}$ is an eigenfunction of $\mathcal{L}_z$ with the eigenvalue $(M + 1)h/2\pi$ unless it is identically zero. But, by (38·15),

$$(\mathcal{L}_x - i\mathcal{L}_y)(\mathcal{L}_x + i\mathcal{L}_y) = \mathcal{L}_x{}^2 + \mathcal{L}_y{}^2 - \mathcal{L}_z\frac{h}{2\pi} = \mathcal{L}^2 - \left(\mathcal{L}_z{}^2 + \frac{h}{2\pi}\mathcal{L}_z\right)$$

$$= \mathcal{L}^2 + \left(\frac{h}{4\pi}\right)^2 - \left(\mathcal{L}_z + \frac{h}{4\pi}\right)^2. \tag{40·18}$$

Hence

$$(\mathcal{L}_x - i\mathcal{L}_y)(\mathcal{L}_x + i\mathcal{L}_y)\psi_{E,L,M} = \left(\frac{h}{2\pi}\right)^2\left[L(L + 1) + \frac{1}{4} - \left(M + \frac{1}{2}\right)^2\right]\psi_{E,L,M}$$

$$= \left(\frac{h}{2\pi}\right)^2\left[\left(L + \frac{1}{2}\right)^2 - \left(M + \frac{1}{2}\right)^2\right]\psi_{E,L,M}. \tag{40·19}$$

[1] P. A. M. Dirac, *P.Q.M.*, 1st ed., p. 90.

Thus $(\mathcal{L}_x + i\mathcal{L}_y)\psi_{E,L,M}$ can vanish only when $M = L$, or $-(L + 1)$. The latter case is impossible, for the relation $\mathcal{L}^2 = \mathcal{L}_x{}^2 + \mathcal{L}_y{}^2 + \mathcal{L}_z{}^2$ applied to a simultaneous eigenfunction of $\mathcal{L}^2$ and $\mathcal{L}_z$ yields

$$\int_\infty \psi_{E,L,M}{}^*(\mathcal{L}^2 - \mathcal{L}_z{}^2)\psi_{E,L,M}d\tau = \left(\frac{h}{2\pi}\right)^2 [L(L + 1) - M^2] \quad \overline{\mathcal{L}_x{}^2 + \mathcal{L}_y{}^2} > 0.$$

(40·20)

We conclude that it is impossible for $|M|$ to be greater than $[L(L + 1)]^{\frac{1}{2}}$ and for M to take on the value $-(L + 1)$. Hence $(\mathcal{L}_x + i\mathcal{L}_y)\psi_{E,L,M}$ is an eigenfunction of $\mathcal{L}_z$ with the eigenvalue $(M + 1)h/2\pi$ unless $M = L$, in which case the function vanishes identically. Similarly

$$(\mathcal{L}_x - i\mathcal{L}_y)\psi_{E,L,M}$$

is an eigenfunction of $\mathcal{L}_z$ with the eigenvalue $(M - 1)h/2\pi$ unless

$$M = -L,$$

in which case it vanishes identically. In short, the existence of the simultaneous eigenfunction $\psi_{E,L,M}$ implies the existence of a set of $(2L + 1)$ such functions: $\psi_{E,L,-L}, \psi_{E,L,-L+1}, \cdots, \psi_{E,L,M}, \cdots, \psi_{E,L,L}$. This means that if $\mathcal{L}^2$ and $\mathcal{L}_z$ commute with the Hamiltonian H (free atom or molecule), an energy level compatible with the quantum number L must have $(2L + 1)$-fold degeneracy and must be compatible with all values of M from $-L$ to $+L$. The manifold made up of the $2L + 1$ simultaneously eigenfunctions of H, $\mathcal{L}^2$, and $\mathcal{L}_z$ with their linear combinations will be called an L *complex*.

In the above discussion L and M were assumed from the beginning to be integers as required by Sec. 34. For future reference, however, it is desirable to note that with one slight ambiguity the same argument can be used to derive the eigenvalues $\mathcal{L}^2$ and $\mathcal{L}_z$ as well as to determine the degeneracy of a pair of mutually compatible eigenvalues of H and $\mathcal{L}^2$.

Let us accordingly consider a set of three dynamical variables α, β, γ which obey the commutation rules

$$[\alpha,\beta] = -\gamma \qquad [\beta,\gamma] = -\alpha \qquad [\gamma,\alpha] = -\beta. \qquad (40\cdot21)$$

Let us denote the typical eigenvalue of γ by the symbol $mh/2\pi$ without making any assumption regarding the values, integral or non-integral, which the parameter m can take on. Let ω^2 denote the operator

$$\omega^2 = \alpha^2 + \beta^2 + \gamma^2. \qquad (40\cdot22)$$

We denote the eigenvalues of ω^2 by $l(l + 1)h^2/4\pi^2$, where l like m is to be treated as a complete unknown.

α, β, γ are assumed to transform class D functions into class D functions. Replacing $\mathcal{L}_x$, $\mathcal{L}_y$, $\mathcal{L}_z$ in the above argument by α, β, γ, respectively, we see that the existence of a simultaneous eigenfunction of H,ω^2,γ, say $\psi_{E,l,m}$ implies the existence of the set: $\psi_{E,l,-l}, \psi_{E,l,-l+1}, \cdots, \psi_{E,l,m}, \cdots, \psi_{E,l,l}$. The values of m form a series, the successive members of which differ by unity and which terminates at the values $\pm l$.

Hence l and m must both be restricted to integral values, or to values which are odd multiples of $\frac{1}{2}$. The operators $\mathfrak{L}_x$, $\mathfrak{L}_y$, $\mathfrak{L}_z$ give a special case in which m and l take on integral values. In the discussion of electron spin (Sec. 61b) we shall see that identifications of the operators α, β, γ, ω^2 exist in which the "quantum numbers" l and m are restricted to half-integral values (*i.e.*, odd multiples of $\frac{1}{2}$).

Having established the fact that the mutual compatibility of the energy level E and the angular-momentum quantum number L implies the existence of an L complex of simultaneous eigenfunctions of H, $\mathfrak{L}^2$, and $\mathfrak{L}_z$, we turn to the question of the additional degeneracy to be expected from the permutation group of symmetry operators. A complete discussion of this question will not be attempted, but one important fact about it will be established.

Let us assume that by applying the operators of this group to the function $\varphi_{E,L,M}^{(1)} = \psi_{E,L,M}$ it is possible to generate $N - 1$ linearly independent new functions, $\varphi_{E,L,M}^{(2)}, \cdots, \varphi_{E,L,M}^{(N)}$. Because every permutation operator commutes normally with H, $\mathfrak{L}^2$, $\mathfrak{L}_z$ each of these functions is a simultaneous eigenfunction of H, $\mathfrak{L}^2$, $\mathfrak{L}_z$ which belongs to E, L, and M. Without loss of generality we can assume that the functions in this set are normalized and mutually orthogonal. By application of the operators $\mathfrak{L}_x \pm i\mathfrak{L}_y$ to each of the functions $\varphi_{E,L,M}^{(k)}$ we can generate a complete corresponding L complex. In this way we obtain a total of $N(2L + 1)$ wave functions for the given pair of values of E and L. All of these wave functions are orthogonal and hence linearly independent. Consider, for example, the pair of functions $\varphi_{E,L,M+1}^{(k)}$, $\varphi_{E,L,M+1}^{(k')}$, where $k \neq k'$:

$$
\begin{aligned}
(\varphi_{E,L,M+1}^{(k)}, \varphi_{E,L,M+1}^{(k')}) &= ((\mathfrak{L}_x + i\mathfrak{L}_y)\varphi_{E,L,M}^{(k)}, (\mathfrak{L}_x + i\mathfrak{L}_y)\varphi_{E,L,M}^{(k')}) \\
&= ((\mathfrak{L}_x^2 + \mathfrak{L}_y^2)\varphi_{E,L,M}^{(k)}, \varphi_{E,L,M}^{(k')}) = ((\mathfrak{L}^2 - \mathfrak{L}_z^2)\varphi_{E,L,M}^{(k)}, \varphi_{E,L,M}^{(k')}) \\
&= [L(L + 1) - M^2]\frac{h^2}{4\pi^2}(\varphi_{E,L,M}^{(k)}, \varphi_{E,L,M}^{(k')}) = 0.
\end{aligned}
$$

The problem of evaluating N is more difficult than that of determining the size of an L complex. Wigner[1] has dealt with the problem by the methods of group theory and Hund[2] has given an alternative procedure which avoids those methods. However, it would be of little use for us to give a general discussion of this problem here since we have not yet introduced the electron-spin coordinates and the Pauli exclusion principle into the theory. In Chap. XIV we shall show how to deal with the question very simply in the light of the exclusion principle and the theory of electron spin.

[1] E. WIGNER, *Zeits. f. Physik* **40**, 492 (1926), **40**, 883 (1927), **43**, 624 (1927).

[2] F. HUND, *Zeits. f. Physik* **43**, 788 (1927).

CHAPTER IX

THE MEASUREMENT OF DYNAMICAL VARIABLES

41. GENERAL THEORY OF MEASUREMENT

41a. Fundamental Characteristics of Measurements.—Having given a purely mathematical definition of a general dynamical variable α, it behooves us to discuss the nature of the process of observing, or measuring, this variable. In classical physics and in quantum physics an observation is basically an operation involving an interaction between the system, or object, under observation, and an observing mechanism. The interaction between the observed system and the observing mechanism terminates in sense perceptions on the part of the observer. In addition to this physical operation there must be a rule, based on definition or theory, for the computation of a number, or numbers, from the sense impressions.

The purpose of a measurement is always to determine some property of the system, whether that property be a fixed one determined by the inherent structure of the system, such as the atomic number of an atom, and defining, or defined by, its Schrödinger equation, or a variable property depending on its *subjective state*. Our present discussion has to do solely with the measurements of the latter kind, *i.e.*, measurements of the quantum analogues of the dynamical variables of classical mechanics. We therefore assume that the structure of the system observed, and its Schrödinger equation, are known *a priori*.

The problem of relating the experimental procedure for the measurement of a dynamical variable with the mathematical definition can be approached from either side. As in the discussion of momentum in Sec. 15 we can assume an experimental procedure and attempt to deduce from it the nature of the corresponding mathematical operator, if any; or, we can start from the mathematical definition of a dynamical variable as a class of operators and attempt to devise an experimental procedure which will measure the eigenvalues. In either case a theory of the interaction between the observed system and the observing mechanism is necessary.

It is convenient to distinguish from the beginning between *individual measurements* designed to attach a specific number to an individual system, and *statistical measurements* designed to determine the possible eigenvalues of an operator and their relative probabilities for an assemblage of identical systems. As previously indicated, the test of quantum-

mechanical theory is to be found in statistical measurements. A contradiction between theory and experiment can be established by means of an individual experiment if it yields a result in conflict with the scheme of theoretically allowed eigenvalues. Complete confirmation of the theory can be obtained, however, only with the aid of a suitable assemblage. Hence statistical experiments are of primary concern.

A statistical measurement designed to test the predictions of wave mechanics should consist ideally of many entirely distinct individual measurements made on the elements of an ideal Gibbsian assemblage of similarly prepared identical systems. Where these systems are of a microscopic character, however, it is not customary to make observations in this ideal laborious fashion. As indicated in Sec. 14b, p. 55, we can then substitute wholesale observations on a concrete assemblage of approximately independent systems for the series of individual observations. Thus an ordinary spectroscopic measurement obtained by a single exposure of a photographic plate gives the statistical distribution of the energies of the photons emitted by a particular light source.

It is to be noted that an attempted individual observation can be either successful or unsuccessful. For example, if one attempted to measure the energy of an individual photon by sending it through a spectroscope, it might be reflected from one of the prism faces and so fail to reach the detecting apparatus. Or, if one attempted to observe the position of an electron by arranging a collision with a second electron, the experiment might fail owing to poor marksmanship. Usually an apparatus for the statistical measurement of a dynamical variable is theoretically capable of giving individual measurements when used in conjunction with individual observed systems, but there may be serious practical difficulties about the latter in the case of atomic systems and the efficiency of such observations (percentage of successful experiments) would frequently be small.

At the outset of this discussion it will be well to emphasize the difference between a scheme of measurement which is satisfactory from the classical standpoint and one which is satisfactory from the quantum-mechanical point of view. In the case of macroscopic mechanical systems the momenta are so large, on the average, that diffraction effects and momentum changes due to observations of position are negligible. For such systems a coordinate and its conjugate momentum are simultaneously measurable, in theory, with an error which can be ignored for all ordinary purposes. It follows that any function of the positional coordinates and their momenta is measurable with a precision which is wholly satisfactory from the point of view of classical dynamics.

In the case of microscopic systems, however, the limitation on the simultaneous measurements of coordinates and momenta inherent in the Heisenberg uncertainty principle is a serious one. Thus the measurement of atomic energies by a calculation based on approximate simultaneous observations of the coordinates and momenta is both theoretically and practically impossible. Happily, however, methods have been developed for the measurement of atomic energies (the spectroscopic and electron-collision methods) which yield consistent results of very

great accuracy. These methods define the energies of atomic systems operationally, and our calculations indicate that the values observed are the eigenvalues of the corresponding Schrödinger Hamiltonian operators. We are thus led to require of a quantum-mechanically satisfactory method of measurement that it shall be capable in principle of yielding results of essentially arbitrary precision independent of limitations due to the Heisenberg uncertainty principle. Such a measurement must not involve the simultaneous determination of the values of conjugate coordinates and momenta. In view of the results already obtained for the energy and momentum we look for the correlation of every measurement of this type with a corresponding Hermitian operator.

41b. Pure States and Mixtures.—We have assumed from the beginning the existence of assemblages of identical systems so prepared that their statistical properties can be described by means of a wave function.[1] Here the statement that an assemblage exists is meant to imply merely that in principle it is possible to prepare such an assemblage. The union of different identically constructed systems into an assemblage is a mental operation rather than a physical one. Hence it is necessarily possible to unite the systems of two or more assemblages into a single superassemblage. Let N_A and N_B denote the numbers of systems in the pure state assemblages A and B. We refer to the ratios $w_A = \dfrac{N_A}{N_A + N_B}$,

$w_B = \dfrac{N_B}{N_A + N_B}$ as the *weights* of A and B in the superassemblage $A + B$. Let W_A, W_B denote the probabilities of an event, or experimental result, for the assemblages A and B, respectively. Let W_{A+B} denote the probability of the same event for the combined assemblage. As the systems in A and B are completely independent, it is necessary that

$$W_{A+B} = w_A W_A + w_b W_B. \tag{41·1}$$

So much follows from the basic rules regarding the calculation of all independent probabilities.

Following von Neumann we designate an assemblage whose statistical properties can be described by a single wave function as a *pure case*, or *pure state*. An assemblage obtained by mixing two or more pure-case assemblages, or having the statistical properties of one so obtained, will be called a *mixed-case assemblage*, or briefly, a *mixture*.

The expectation value of a dynamical variable α for a pure-state assemblage with wave function Ψ is $\bar{\alpha} = (\alpha\Psi,\Psi)$. In the case of a mixture of pure-case subassemblages with the wave functions Ψ_1, Ψ_2, $\cdots \Psi_j$, $\cdots$ and the weights w_1, w_2, $\cdots w_j$, $\cdots$ we have

$$\bar{\alpha} = \sum_j w_j(\alpha\Psi_j,\Psi_j).$$

[1] *Cf.* pp. 52, 70.

The question now arises whether a mixture of two or more pure-case subassemblages is equivalent to a single pure-case assemblage or is actually something more general.

To settle this question we assume that the answer is affirmative and prove a contradiction. Let the wave functions $\Psi_1, \Psi_2, \cdots \Psi_j, \cdots \Psi_n$ of the components of the above specified mixture be simultaneous eigenfunctions of a complete normal set of independent commuting dynamical variables $\alpha_1, \alpha_2, \cdots$. Let $a^{(j)}$ denote the set of eigenvalues of the α's which belongs to Ψ_j. These sets are assumed all different. The probability of $a^{(j)}$ for the mixture is w_j. But if the mixture is equivalent to a pure-case assemblage, the corresponding wave function must be a linear combination of the wave functions of the subassemblages, say,

$$\Psi = \sum_{j=1}^{n} c_j \Psi_j. \tag{41.2}$$

The probability of the set of eigenvalues $a^{(j)}$ computed from Ψ is $|c_j|^2$. Consistency demands that $|c_j|^2 = w_j$. In order that Ψ shall also represent the statistical properties of the superassemblage with respect to the positional coordinates it is necessary that

$$|\Psi|^2 = \sum_{j=1}^{n} w_j |\Psi_j|^2, \tag{41.3}$$

whereas (41.2) implies that $|\Psi|^2 = |\sum_{j=1}^{n} w_j^{1/2} \Psi_j|^2$. These equations are in general inconsistent. Hence we conclude that in general a mixture of pure cases is not a pure case. An important exception occurs when we mix several pure-case assemblages whose wave functions are multiples of one another. It is immediately evident that in a case of this kind, Eqs. (41.2) and (41.3) involve no contradiction. Hence such assemblages can always be united into a single superassemblage which is a pure state with a wave function which is a multiple of the wave function of each of the subassemblages.

When all we know of the state of a system is that it belongs to a mixture of pure-state assemblages, we shall speak of it as being in a mixed state.

The difficulty which makes it impossible to replace a mixture of non-identical pure-state assemblages with a single equivalent pure-state assemblage also prevents the resolution of a pure-state assemblage into a mixture of non-identical pure-state subassemblages. von Neumann has treated these questions at length.[1] The important fact for us to note here is that an assemblage of similarly prepared identical systems must be assumed to be a mixture rather than a pure state unless the

[1] VON NEUMANN, *M.G.Q.*, IV, 1–3.

method of preparation has been such as to insure that no system shall be admitted to the final assemblage which is not in the same state as every other system. The methods used in preparing such a pure-state assemblage have been indicated in Sec. 15f, p. 70.

41c. Postulates Regarding Retrospective and Predictive Measurements.—There are two different aspects of measurements or observations which it is convenient to distinguish at this point. These are the backward-looking aspect and the forward-looking aspect. Let us consider them first of all from the classical standpoint. In classical theory there is no reason why the measurement of a dynamical variable need alter its value or change the state of the system under observation. Hence it is customary to assume that classical measurements give the values of the quantities observed both before and after the measurement. Actually, however, there are plenty of classical processes which we can hardly fail to classify as measurements in which the state of the system is not preserved. A chemical analysis does not leave the material under investigation in its original state but gives properties of the sample analyzed before the measurement. It may result in the preparation of samples of pure materials and hence yields data concerning the properties of these end products. The information regarding the future is thus of an entirely different character from that regarding the past. Moreover the backward-looking and forward-looking aspects of a measurement can be completely dissociated. It is possible to make backward-looking measurements which change the state of the system under observation and yield no definite information regarding its final state. We can also make forward-looking measurements which have no backward-looking aspect. The former is illustrated by a measurement of momentum following the method of Sec. 15, but making the final observation of position by a violent impact. The latter is illustrated by methods of factory production in which a standardized product is turned out by means of a punch or templet.

From the standpoint of quantum mechanics the distinction between the backward-looking and forward-looking aspects of measurement is more important than from the standpoint of classical theory because the interaction of the observing mechanism with the observed system, ignored in the classical theory, prevents the design of measurements which always leave the system in exactly the same state after the observation as before. It is sometimes possible to carry out a measurement so that if the system observed is initially in an eigenstate of the dynamical variable α which is to be measured, its wave function, and hence its state, are the same after the measurement as before. With the same experimental arrangement, however, the wave function will be changed by the measurement if its initial form is not that of an eigenfunction. Thus in the case of any measurement of a microscopic system capable of

exhibiting quantum-mechanical properties, the information obtained regarding the initial state is the same as that regarding the final state, only if the initial state itself, as well as the experimental procedure, is properly chosen. We accordingly label a measurement as *retrospective* or *predictive* according as it yields information about the immediately preceding or immediately succeeding state of the measured system. Of course, it is possible for a measurement to be both retrospective and predictive.

After the above preliminary discussion we are in a position to lay down certain theoretical requirements, or postulates, regarding the measurement of a dynamical variable α defined as in Sec. 36.

a. *A successful, exact, individual measurement of* α, *whether retrospective, or predictive, must always yield as its numerical result one of the eigenvalues of* α.

b. *If an exact individual predictive measurement is followed immediately by a successful exact individual retrospective measurement, the two results must agree.*

c. *The probability of any eigenvalue* α' *or any range of eigenvalues* $d\alpha'$ *in an individual retrospective measurement of* α *for a system initially belonging to a pure state assemblage must be in harmony with the* Q *formulas* (36·77) *and* (36·78).

In the case of a mixed type of assemblage the probability will then be given automatically by the appropriate generalization of (41·1).

In conformity with **c** the theoretical calculation of the probability of an elementary range of eigenvalues $d\alpha'$ for a pure-state assemblage is to be carried out as follows: First one must resolve the Cartesian wave function ψ_x into a series-integral (spectrum) whose elements are eigenfunctions of α. One next forms the partial sum (or integral), say $(\Psi_x)_{d\alpha'}$, of those elements of this spectrum which belong to the range $d\alpha'$ and identifies the desired probability $Q(\alpha')d\alpha'$ with the scalar product, or integrated intensity, $((\Psi_x)_{d\alpha'}, (\Psi_x)_{d\alpha'})$ of this partial wave function.[1]

---

[1] According to the scheme of Eq. (36·77) one should choose a complete set of dynamical variables q_1, q_2, $\cdots$ of which α is a member and compute $Q(\alpha')d\alpha'$ from the probability amplitude

$$\Psi_q(q') = T^{\binom{q}{x}}\Psi_x(x').$$

$\Psi_q(q')$ can be regarded as the sum of orthogonal elements $(\Psi_q)_{d\alpha'}$, each of which is equal to Ψ_q in the corresponding range of α eigenvalues and vanishes outside that range. The transformation of these elementary functions back into Cartesian-coordinate space, according to the rule

$$(\Psi_x)_{d\alpha'} = (T^{-1})^{\binom{x}{q}}(\Psi_q)_{d\alpha'},$$

gives a set of exact or approximate eigenfunctions of α in Cartesian space whose sum is Ψ_x. Owing to the unitary character of $T^{\binom{q}{x}}$ the integrated intensity of $(\Psi_x)_{d\alpha'}$ in Cartesian space is equal to the integrated intensity of $(\Psi_q)_{d\alpha'}$ in q' space, which is evidently equal in turn to the value of $Q(\alpha')d\alpha'$ worked out from Eq. (36·77).

As previously indicated, any procedure for dealing simultaneously with all the elements of an assemblage of sensibly independent systems, which yields the eigenvalues of α and their probabilities, is to be reckoned a valid statistical measurement, even though it is not, strictly speaking, a compound of many individual measurements.

It follows from **b** that an assemblage of similar systems correlated with a common eigenvalue α' as a result of a predictive measurement must be of such a character that a statistical retrospective measurement will show the probability of α' to be unity. Hence this assemblage must be a mixture of pure-state subassemblages each of which has a wave function which is an eigenfunction of α with the eigenvalue α'. In special cases the mixture can reduce to a pure state. This will happen, for instance, if the eigenvalue α' under consideration is nondegenerate, so that all the corresponding eigenfunctions are linearly dependent and physically equivalent.

In the case of a pure-state assemblage prepared by a predictive measurement the wave function must be "physically admissible," as it is created by physical manipulation rather than by mathematical definition. This means that it is impossible to carry out exact predictive measurements of dynamical variables which do not have class D eigenfunctions. We are accustomed, however, to idealize measuring instruments by treating them on a classical basis and ascribing to them discontinuous properties, such as smooth surfaces and edges, which from a microscopic point of view they do not possess. If such idealizations are allowed in the theory of a measurement, they produce discontinuities in the wave functions of assemblages prepared by predictive measurement. Hence the assumption that the wave functions of pure-state assemblages prepared by predictive measurements must be of class D is very inconvenient. We shall therefore introduce a milder restriction in the form of a fourth postulate, which implies the admission of idealized measurements along with those which are actually possible.

d. *The eigenfunctions of pure-state assemblages prepared by predictive measurements, or of the pure-state subassemblages which make up a mixed assemblage prepared in this way, are all quadratically integrable.*

As the eigenfunctions of continuous-spectrum eigenvalues are not quadratically integrable, we infer from **d** that *exact predictive measurements of continuous-spectrum eigenvalues are fundamentally impossible.*[1] Experience shows that exact retrospective measurement of such eigenvalues is equally impossible, although this fact is not an immediate consequence of our present postulates. In these cases where exact measurements are not possible, we assume that inexact measurements,

---

[1] For many purposes it is convenient to reckon discrete eigenvalues as exactly measurable, but the above statement should not be understood to imply that such exact measurements are strictly speaking realizable.

which conform to **a** and **d** and satisfy **b** and **c** to an arbitrarily high degree of approximation, are to take their place.

The reader will now inquire by what means we are to make connections between the above requirements **a, b, c** and concrete experimental methods. In discussing this general problem it will be of advantage to examine first the case of positional measurements. They occupy a special place in the development of the theory because of our initial physical interpretation of $|\Psi|^2$ as probability density in coordinate space. In fact our present postulates when applied to positional measurements reduce to an amplification of assumption **b** of Sec. 14*a*.

The interpretation of $|\Psi|^2$ as probability density has operational meaning only if we have in mind one or more mutually consistent definite methods of observing the spatial configurations of the systems described by Ψ. Of course we do have such methods for observing the positions and orientations of large-scale bodies. They depend primarily on optical coincidences of one kind or another, although the senses of touch and hearing could be called into play for observations of this kind. Not to become too discursive, it will be assumed that any of the classical methods for the measurement of configuration is valid in the domain of quantum mechanics if due regard is paid to the finite resolving power of optical instruments and to the diffraction of matter corpuscles in methods of observation by collision. There is a difficulty in principle about the observation of unique atomic configurations owing to the fact that they contain many identical particles (electrons) between which we are unable to distinguish. For the moment, however, we shall ignore this difficulty and proceed with the discussion as if all electrons and other like particles were provided with legible number plates.

Since all values of positional coordinates are eigenvalues of the corresponding operators, it is clear that postulate **a** is automatically satisfied by any of the necessarily approximate positional observations in the sense that it yields a short range of possible values from which any one can be arbitrarily chosen as the result of the measurement. In connection with postulate **b** we may assume that any observation which from a classical point of view gives the position at the moment when the observing mechanism *begins* to interact with the system is a retrospective observation, while one which gives the position at the moment when the observing system *ceases* to interact is a predictive observation. In general an observation is both retrospective and predictive if the time of interaction between the observed system and the observing mechanism is small, and the variable measured is conserved from a classical point of view during the interaction. As the positional coordinates of a free dynamical system are not constant in time, we have formulated **b** as a statement of agreement between a predictive observation and a retrospective one performed *immediately afterwards*. When so

formulated, the validity of **b** for positional observations within the limit of available precision is a matter of common experience.

Postulate **c** is the fundamental postulate of wave mechanics. It can be tested only in conjunction with such other assumptions as yield an initial Ψ appropriate to the state in question. We take its validity for atomic systems to be a matter of common experience supported by a considerable variety of experiments in which the spatial distributions of electrons and atoms are found to agree with wave-mechanical calculations.

Consider next the applicability of **a, b, c** to the measurement of dynamical variables which are not spatial coordinates. There is an important class of measurements for which the validity of these postulates is a direct consequence of their validity for observations of configuration, *viz.*, those measurements which involve an interaction with an observing mechanism (in limiting cases this step is omitted) followed by a positional observation carried out either on the observed system, or on a part of the observing mechanism. In such cases the correctness of the postulates can be proved or disproved from the hypotheses regarding $|\Psi|^2$ by an appropriate quantum-mechanical analysis of the assumed experimental procedure. We have already done this for linear-momentum measurements in Sec. 15. In another class of measurements the interaction between radiation and matter plays a fundamental part, and the theory of the experiment becomes a special case of quantum electrodynamics which involves postulates not included in the general theory here developed. Even for measurements of this type, however, *ad hoc* hypotheses should be unnecessary, the validity of the postulates **a, b,** and **c** being a consequence of the basic assumptions of the theory.

These considerations lead us to the formulation of an additional postulate:

e. *The statistical distribution of results in the case of any type of experiment performed upon the members of a pure-state assemblage is calculable in principle from the corresponding wave function.*

The experiments under consideration here include many which are not statistical measurements of a dynamical variable. Measurements of the scattering of electrons by atoms, for example, are not simply measurements of the eigenvalues of a dynamical variable but determine the change in the statistical distribution of momenta of an assemblage of free electrons brought about by interaction with atoms. The postulate means that a pure-state assemblage has no intrinsic properties not implicit in its wave function.

41d. The Reduction of the Wave Packet.[1]—We have already shown (p. 324) that in consequence of postulates **a, b,** and **c** a statistical predictive measurement of a dynamical variable α must prepare an assemblage which is a mixture of subassemblages each of which is an eigenfunction

---

[1] *Cf.* Secs. 14*b*, 15*f*, 16 (p. 75), and 19*d*.

of α with a common eigenvalue α'. Therefore, in general, there must be a discontinuous change in the wave function or mixture of wave functions associated with a system when one of its dynamical variables is measured. As this change has confused some physicists it will be well to analyze it carefully. If the observing mechanism, owing to its relatively large mass, can be treated on a classical basis, the operation of measurement converts each initial wave function into a definite corresponding final wave function, *i.e.*, it converts pure states into pure states. The scope of the following discussion is limited to measurements of this simplest variety. In Sec. 42 we shall deal briefly with measurements of a more general type.

When a measurement converts pure states into pure states the discontinuous change in the wave function which accompanies a measurement is called the *reduction of the wave packet*. The necessity for such a change has been noted in the earlier sections of this book (*cf.* Sec. 15*f*). It is clearly unavoidable if we accept the notion that $|\Psi|^2 d\tau$ is a measure of the probability of configurations which lie within the element $d\tau$ of configuration space, and if we assume the validity of b for measurements of position. To put the argument in its most elementary form consider a single-particle problem with a wave function $\Psi(x,y,z,t)$ spread out over a large effective volume V at the instant t_0. Let a positional measurement made at t_0 locate the particle within a relatively small volume element δV. If this measurement is properly carried out it will leave the particle in the element δV at t_0, although it will inevitably alter the momentum. Thus a wave function which is to determine the probabilities of different positions for *future* times must approach zero outside δV as t decreases to t_0. Otherwise there could not be that agreement between successive observations of position which creates the continuity of our experiences with large-scale objects.

The assumption of such a discontinuous change in the wave function is paradoxical, however, in view of the continuity of Ψ as a function of t demanded by the fundamental Schrödinger equation (7·3). This paradox becomes acute if one attempts to interpret wave functions as in some sense physical realities. It is possible, for example, to determine the position of a particle A by an indirect method in which it is allowed to collide with a second particle B whose "orbit" before and after the collision is measured. In this case the wave function of the combined system $A + B$ is reduced when the scattered particle B is located after the collision. This reduction locates the particle A at the moment of the collision and also determines its momentum (within the limits allowed by the Heisenberg uncertainty principle). Thus an observation made on B which does not involve any disturbance of A can alter completely our expectation of the future behavior of the latter and thus completely change the type of wave function which must be associated with it

if we are to treat it as an independent system.[1] Clearly this fact should reinforce our previous conclusion that the wave function is merely a subjective computational tool and not in any sense a description of objective reality.

On the other hand the paradox is at once resolved if we adhere strictly to the previously formulated interpretation of Ψ as a mathematical specification of the properties of an assemblage of similarly prepared, identical, independent systems. Whenever a statistical predictive measurement is carried out we have to do with a *selective process* in which
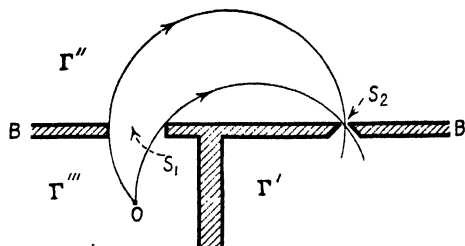
Fig. 18.—The magnetic-deflection method of measuring energies and momenta.

the systems of the original assemblage are divided into two or more sub-assemblages each associated with a value, or range of values, of the dynamical variable measured. The discontinuous change in the wave function of a system which accompanies the measurement of a variable α is the reflection of a mental process in which we transfer the system under consideration from an initial assemblage A to a subassemblage $A_{\alpha'}$, consisting of those members of A which have given the measured value α' of α, all other members of A having been rejected. There is evidently no contradiction between the discontinuous change in the wave function of a system which accompanies this mental transfer and the continuity required by Eq. (7·2) for the wave function of a fixed assemblage of systems subject to definite external forces.

The situation will be clarified by an example. Consider a measurement of electron energies by the magnetic-deflection method. Let O in Fig. 18 be the source of the stream of electrons measured, let S_1 be a slit and BB the focal plane. Diffraction effects at S_1 are small with the usual energies and slit widths so that a classical theory is adequate to tell us that a uniform magnetic field perpendicular to the plane of the paper will bring electrons with any given energy to an approximate focus in the plane BB provided that BB is perpendicular to the line drawn from

---

[1] This general type of observation has been the subject of much recent discussion. (*Cf.* Einstein, Podolsky, and Rosen, "Can Quantum-mechanical Description of Physical Reality Be Considered Complete," *Phys. Rev.* **47**, 777 (1935); N. Bohr, *Phys. Rev.* **48**, 696 (1935); E. Schrödinger, *Naturwissenschaften* **23**, 807, 823, 844 (1935), *Proc. Cambridge Phil. Soc.* **31**, 555 (1935); W. H. Furry, *Phys. Rev.* **49**, 493 (1936).

O to the mid-point of S_1.   A photographic plate located on this plane provides a possible basis for a retrospective statistical measurement of the electron energies.   Of course, it is necessary for the success of the experiment that the current drawn from O shall be small enough so that the Coulomb forces between the electrons are negligible.   Only in that case can we consider the stream to be the equivalent of a Gibbsian assemblage of independent electrons.

To convert this retrospective measurement into a predictive one a diaphragm cut by one or more slits S_2 can be substituted for the photographic plate.   The electrons which pass through any of the slits then form a subassemblage with energies within a small range ΔE determined by the slit widths and resolving power.   This modification of the experimental arrangement is analogous to the conversion of an optical spectrograph into a monochromator.

A first step in the quantum-mechanical analysis of the experiment is to note that in general the assemblage of electrons leaving the source O will not be of the pure-state type but a mixture of such states.   We are at liberty, however, to deal individually with the component pure-state subassemblages.   Let us concentrate our attention on one of these components, say γ, whose three-dimensional wave function we designate as $\Psi_\gamma(x,y,z,t)$.   We assume that for our present purpose it is legitimate to idealize the diaphragms and other barriers to the motion of free electrons, treating them as infinite potential barriers which reflect all incident Ψ waves.[1]   The interaction of the diaphragm with the electrons then produces no changes in Ψ_γ which are discontinuous in time, but it does divide it into three parts: a portion $\Psi_{\gamma'}$ which passes through S_1 and S_2; a portion $\Psi_{\gamma''}$ which passes through S_1 but strikes the upper side of the diaphragm after magnetic deflection; and a portion $\Psi_{\gamma'''}$ which fails to pass S_1.   (For simplicity we assume a single slit S_2 in the path of the downward-moving electrons.)   If the experimenter desires to separate the electrons which pass through each slit from those which are blocked off, and to keep them separate, he will interpose barriers which actually prevent any overlapping of $\Psi_{\gamma'}$, $\Psi_{\gamma''}$, $\Psi_{\gamma'''}$ once these three portions of Ψ_γ have been separated.   From a classical point of view this procedure would divide the original assemblage γ into three distinct subassemblages of electrons γ', γ'', γ''' each located, after the closing of the slits, in a different vessel.   For convenience we designate by Γ', Γ'', Γ''' the three vessels, or portions of space, in which $\Psi_{\gamma'}$, $\Psi_{\gamma''}$, $\Psi_{\gamma'''}$ are laid out.   If there is no condensation on the walls, the number of electrons in each of these vessels would be constant in classical theory from the moment of closing the slits onward.   Actually it is operationally meaningless to say that the number of electrons in each vessel is constant

---

[1] A mathematical model with a complex potential energy would provide for absorption of some electrons at the walls of containing vessels, if desired.

between the time when the slits are closed and the first time—if any—when the wave functions are disturbed by a counting operation. However, the probable number of electrons in each vessel is constant in time during this period according to wave mechanics, being equal to $N\Psi_{\gamma'}$ [*i.e.*, $(\Psi_{\gamma'}, \Psi_{\gamma'})$], $N\Psi_{\gamma''}$, or $N\Psi_{\gamma'''}$ as the case may be.

We now desire to show that the electrons in Γ' which come from the primary assemblage γ may be regarded as forming a pure-state assemblage with the normalized wave function $\Psi_{\gamma'}/(N\Psi_{\gamma})^{\frac{1}{2}}$. (To be sure we do not know *a priori* the exact number of electrons in this assemblage as we should in the case of an ordinary Gibbsian assemblage. We know only that the probable value of the ratio of the number of electrons in the volume Γ' to the total number in the assemblage γ is $N\Psi_{\gamma'}$. This restriction is of no fundamental importance, however, since in practice we never know *a priori* the number of elements in one of the concrete assemblages of approximately independent systems with which we are forced to deal in testing the predictions of wave mechanics for atomic problems.) Consider first the result of a retrospective positional observation made upon the electrons in Γ'. As these electrons form part of the pure-state assemblage γ, the ratio of the probable number of electrons in any volume element dV of Γ' to the whole number of electrons in γ is $|\Psi_{\gamma}|^2 dV$, or $|\Psi_{\gamma'}|^2 dV$. Hence the probable value of the ratio of the number of electrons in dV to the number in all Γ' coming from γ is $|\Psi_{\gamma'}|^2 dV/N\Psi_{\gamma'}$. This proves that, so far as positional measurements are concerned, the electrons in Γ' which come from γ are equivalent to a pure-state assemblage γ' with the normalized wave function $\Psi_{\gamma'}/(N\Psi_{\gamma'})^{\frac{1}{2}}$. But from the discussion in Sec. 41c, p. 326, we learn that the possibility of using a given wave function Ψ to predict the result of an arbitrary statistical measurement is to be regarded as a consequence of the assumption that $|\Psi|^2$ gives the probability density in configuration space. Hence it must be possible to predict the result of any type of statistical measurement applied to the electrons in Γ' by treating these electrons as a pure-state assemblage with the wave function $\Psi_{\gamma'}/(N\Psi_{\gamma'})^{\frac{1}{2}}$.

One possible objection to the above reasoning lies in the fact that we have not allowed for the necessity of verifying in the case of an individual electron that after the slits are closed it actually has turned up in Γ', or Γ'', or Γ''', as the case may be. If we were dealing with an ideal Gibbsian assemblage and carrying through predictive measurements of energy, one at a time, we should not consider our work finished until such a verification had been made in each case. This verification might be made in such fashion as to completely spoil the wave function associated with the future of the electron. Can it also be made in such fashion as to leave that function unaffected? We believe the answer to be affirmative, for it will suffice if we wish to be sure that an individual electron is in Γ',

to search thoroughly in Γ″ and Γ‴ and to verify that it is in neither of these vessels. Such a search—however futile in the case of a collar button—provides a conceptual means for securing the desired information without disturbing the electron or changing in any way the expectation regarding its future behavior.

So far we have ignored all the electrons emanating from the source except those in the pure-state subassemblage γ. It is clear, however, that every pure state σ contained within the original mixture will be split exactly like γ. Hence the complete assemblage of electrons passing through the slit S_2 will behave like a mixture of subassemblages each of which has a wave function derived from the corresponding initial Ψ_σ by a process of division and renormalization like that described above.[1]

The experiment we have analyzed is a very special one, but it will suffice, perhaps, to show that the assumed discontinuous change in the wave function associated with a system when one of its dynamical variables is measured does not violate the rule of continuity implicit in the Schrödinger equation (7·3).

41e. Classical Orbits and Wave Packets.—An interesting point to be noted in connection with the reduction of the wave packet is that the mere interaction of the observing mechanism with the observed system does not make it necessary to reduce the packet. On the contrary, this interaction produces a continuous change. If the packet is to be reduced, the interaction must have produced knowledge in the brain of the observer. If the observer forgets the result of his observation, or loses his notebook, the packet is not reduced. We are again led to emphasize the fact that the wave function of a pure-state assemblage is merely a mathematical tool for computing from all previous observations what the relative probabilities are for different results when we make our next observation.

This interpretation solves the paradox of the fact that, although every wave packet grows in volume indefinitely with time, the motion of any large-scale body seems to follow a sharply defined classical orbit. The

---

[1] In general the wave functions of the different components of the mixed assemblage in Γ′ obtained by the process of predictive measurement will be much more nearly alike than the wave functions Ψ_σ of the original mixture. Each of the new functions has been obtained from its parent by a common process of trimming which restricts the energy range in each case to ΔE and makes a similar reduction in the range of values of the vector momentum. If the measurement is performed in a manner which yields a good resolving power (i.e., with a weak magnetic field, large radius of curvature, and slit widths of the order of the wave length $\lambda = h/p$) the reduced wave functions will be so similar that for most purposes their differences will be negligible. Under these circumstances the mixture in Γ′ will behave like a pure-state assemblage and can be treated as such. In fact the only way in which a pure-state assemblage can be prepared is by starting with a mixture and making such observations as are needed to reduce the wave functions of all its elements to a common form.

experimental fact behind this appearance of continuity is the following. If we make an observation of the position and momentum of a macroscopic body moving under the influence of known forces, and compute correctly from the data by classical methods the position which the body should occupy at some later time t, making due allowance for experimental error, we always find that the measured position at the time t lies within the range permitted by our calculations.[1] Hence we may consider that the paradox is removed if we can show that in such cases the range of uncertainty at the time t allowed by a wave-mechanical calculation is essentially the same as for a classical calculation.

For simplicity let us consider the case of a particle (*i.e.*, the center of gravity of any body) moving in one dimension under no forces. We shall suppose that our observations have given us a probable initial position $\bar{x}_0$, a probable momentum $\bar{p}$ (this will be constant for a free particle), and corresponding values of the root-mean-square deviations of the actual values of these quantities from their estimated values. We denote these root-mean-square deviations by Δx_0 and Δp, respectively. From a *classical point of view* the values of $\bar{x}_0$, Δx_0, $\bar{p}$, Δp, together with the assumption of the Gaussian law of error, serve to define a probability distribution in the two-dimensional classical phase space, and so to fix a Gibbsian assemblage of systems having the property that the probability of any range of values of x or p for the individual system observed is the same as the probability of the same range for the assemblage. Thus, even classically, the prediction of the future states of an observed system is reduced to statistical form when the experimental errors are taken into account.

For each element of this hypothetical classical assemblage

$$x_t = x_0 + \frac{pt}{m}.$$

Forming averages of each side of this equation over all elements of the assemblage we obtain

$$\bar{x}_t = \bar{x}_0 + \frac{t}{m}\bar{p}. \tag{41.4}$$

To obtain the root-mean-square deviation of x_t from $\bar{x}_t$, we express $(x_t - \bar{x}_t)^2$ in terms of x_0, $\bar{x}_0$, p, $\bar{p}$, t, and average over the assemblage. The resulting expression for $(\Delta x_t)^2$ is

$$(\Delta x_t)^2 = \overline{(x_t - \bar{x}_t)^2} = (\Delta x_0)^2 + \frac{2t}{m}(\overline{px_0} - \bar{p}\bar{x}_0) + \frac{t^2}{m^2}(\Delta p)^2. \tag{41.5}$$

Thus the classical uncertainty in the position of the particle increases with time by an amount which depends on the uncertainty in the momentum.

[1] The continuous motion of the image of a large-scale body on the retina of one's eye may be regarded as a limiting case of the above general proposition.

From the *quantum-mechanical point of view* our expectation regarding the probability of different ranges of values of x_t is again to be computed with the aid of a suitably defined Gibbsian assemblage. The only difference is that the individual systems of the assemblage do not have definite pairs of values of x and p at any time. In this case the assemblage to be dealt with is a mixture of pure-state subassemblages, each of which has a definite probability w_k and a definite wave function Ψ_k. The expected position of the particle at the time t is to be obtained by computing from each of the constituent wave functions a corresponding mean value of x_t and then averaging these over the different subassemblages. Thus,

$$\bar{x}_t = \sum_k w_k \int x |\Psi_k(x,t)|^2 dx. \tag{41.6}$$

Similarly,

$$\overline{(x_t - \bar{x}_t)^2} = \sum_k w_k \int (x - \bar{x}_t)^2 |\Psi_k(x,t)|^2 dx. \tag{41.7}$$

From Eq. (33·18) it follows that Eq. (41·5) applies in quantum theory as well as in classical theory provided that the indicated averages are carried out according to the rule of Eqs. (41·6) and (41·7), and provided $\overline{px_0}$ is replaced by its quantum-mechanical equivalent $\dfrac{\overline{px_0} + \overline{x_0 p}}{2}$ to take into account the effects of non-commutation.[1] But whether we use classical theory or quantum theory the values of Δx_0 and Δp are to be taken from the same experimental data. Finally the value of $\overline{px_0} - \bar{p}\,\overline{x_0}$ to be used in either case must depend on an analysis of the experimental procedure actually used in measuring the initial position and momentum. If the measurements of position and momentum are entirely independent, this quantity should be set equal to zero, but, if the measurement of position at $t = 0$ is used in the calculation of momentum, this will not

---

[1] The calculation can be carried out as follows: Let $A_k[x]$ denote the expectation value of a quantity x in the state characterized by Ψ_k. Thus,

$$\bar{x} = \sum_k w_k A_k[x] = \sum_k w_k A_k[x_0] + \frac{t}{m} \sum_k w_k A_k[p]$$

$$= \bar{x}_0 + \frac{t}{m}\bar{p};$$

$$\overline{(x - \bar{x})^2} = \sum_k w_k A_k[(x - \bar{x})^2] = \sum_k w_k A_k[x^2] - \bar{x}^2 = \overline{x^2} - \bar{x}^2$$

$$= \sum_k w_k \left\{ A_k[x_0^2] + \frac{t}{m} A_k[px_0 + x_0 p] + \frac{t^2}{m^2} A_k[p^2] \right\} - \bar{x}^2$$

$$\cong \overline{x_0^2} + \frac{2t}{m}\overline{px_0} + \frac{t^2}{m^2}\overline{p^2} - \bar{x}_0^2 - \frac{2t}{m}\bar{x}_0\bar{p} - \frac{t^2}{m^2}\bar{p}^2$$

$$\cong \overline{(x_0 - \bar{x}_0)^2} + \frac{2t}{m}(\overline{px_0} - \bar{p}\bar{x}_0) + \frac{t^2}{m^2}\overline{(p - \bar{p})^2}.$$

be true. In either case, however, the uncertainty in x_t allowed by classical and quantum theories under given experimental conditions must be the same.

Although the above discussion is based on a very simple special case it does show that, so long as we regard wave functions as merely tools for the calculation from present measurements of the results to be expected from future measurements, there is no real contradiction between the growth of a wave packet in time and the precision with which large-scale bodies are found by experiment to follow classical trajectories.

42. MORE ABOUT MEASUREMENTS

42a. Conjugate Variables and Measurements.—The impossibility of making simultaneous exact measurements of the Cartesian coordinates and conjugate components of linear momentum has already been discussed in Secs. 16 and 33b. An exact simultaneous retrospective or predictive measurement of two dynamical variables will be possible only if they are "mutually compatible" or have a complete system of simultaneous eigenfunctions. As we saw in Sec. 37c, the mutual compatibility of two variables implies their commutability, so that we have the general rule that only commuting operators are simultaneously measurable. The rule regarding conjugate variables is a special case of this more general principle.

One can show that the predictive measurement of any dynamical variable α has a tendency to destroy preexisting knowledge regarding other variables which do not commute with α. This destruction becomes complete in the case of an exact measurement of one member of a pair of canonically conjugate variables. These are pairs of variables either one of which may be regarded as a coordinate, while the other, or its negative, plays the role of corresponding momentum. The exact definition is given in Sec. 39a. Here it suffices to note that in consequence of the definition, if q_k is one of the coordinates of a legitimate coordinate system $q_1, q_2, \cdots, q_\lambda$, and if p_k is its conjugate momentum, the eigenfunctions of p_k in q' space have the form

$$(q'|p_k', \cdots) = e^{\frac{2\pi i}{h}p_k'q_k'}u(q'),$$

where $u(q')$ is independent of q_k'. Hence the probability function $Q(q_k')$ of Eq. (36·78) becomes

$$Q(q_k') = \overline{\sum_{q' \neq q_k'}} |(q'|p_k', \cdots)|^2 = \overline{\sum_{q' \neq q_k'}} |u(q')|^2,$$

in the case of such an eigenfunction of p_k. As this is independent of q_k', it follows that if p_k' is exactly known, all eigenvalues of q_k become equally probable.

42b. Impossibility of Measurements Which Imply Distinction between Particles of Same Species.—The hypothesis is sometimes made that to every classical dynamical variable there corresponds a uniquely defined quantum dynamical variable. This hypothesis we hold incorrect inasmuch as there are classical variables, such as the radial momentum in a three-dimensional problem, which definitely have no analogues in the class of true quantum dynamical variables.[1]

One would also like to believe that whenever we have found a class of operators which conforms to the mathematical requirements laid down for a dynamical variable in quantum mechanics, there exists at least a conceptual experimental method for measuring the variable exactly. Here again, however, the attractive simple hypothesis breaks down. There is at least one large class of exceptions, and there may be others outside this class.

We suppose matter to be composed of a number of species of elementary particles, the members of each species being exactly alike. As we noted in Sec. 40d it is essentially impossible to distinguish experimentally between two particles of the same species, be they electrons, protons, neutrons, or what not. Therefore the best we can actually do, even in principle, in the way of determining the configuration of an atomic system consisting of, say, a nucleus, which we treat as a single heavy particle, and f electrons, is to determine the coordinates of the nucleus and f sets of coordinates for the f electrons.[2] It is not possible to distinguish between the $f!$ different configurations which go with the $f!$ different correlations of the f electrons with the f observed positions. Hence such an observation gives us a function of the configuration which is the same for all configurations derived from any one by a permutation of the coordinates—in other words, a symmetric function of these coordinates.

Our inability to measure dynamical variables which are not symmetric functions of the coordinates and momenta of like particles implies an inability to prepare pure subjective states whose wave functions are eigenfunctions of nonsymmetric variables. In fact the preparation of a system in a state whose wave function is not in some sense symmetric with respect to the coordinates of like particles must be excluded as

---

[1] In Sec. 39a it was proved that the Hermitian operator which we should expect to play the part of radial momentum does not have a complete set of eigenfunctions and is not a true dynamical variable in quantum mechanics. From the experimental point of view we find ourselves in equal difficulty. Since the radial momentum is not constant, classically, even for a free particle, unless it happens to have zero angular momentum, we see that the scheme for the exact measurement of p_x, p_y, p_z used in Sec. 15 is *not* applicable to the radial momentum. Hence we can be sure that our mathematical difficulties with the radial-momentum operator are not due to an improper choice of that operator.

[2] This might be done with a spray of high-intensity hard γ rays.

implying an ability to distinguish between such particles. Hence we may properly add a corresponding restriction to the manifold of physically admissible wave functions hitherto identified with class D. Evidently this restriction can be put in the following form: *All the physical predictions which flow out of any physically admissible wave function must be unchanged by the application to that function of an arbitrary permutation of the coordinates of like particles.* Otherwise we could make tests to check which arrangement of the coordinates is right and so acquire information inconsistent with the identity of the particles concerned.

Of course the above rule can hold only if *all* the coordinates of each particle appear as arguments in the wave function, and if *all* are subject to permutation. Actually it is necessary (*cf.* Sec. 61a) to add to the three positional coordinates x_k, y_k, z_k of each electron and proton, a fourth "spin" coordinate σ_k, which we have not hitherto considered. Hence the wave functions dealt with up to this point are "incomplete" in that the spin coordinates are suppressed, and the rule we have formulated is not applicable to them. *In the remainder of Sec. 42b the ψ functions discussed are assumed to be of the complete type to which the permutation restriction does apply.*

Recollecting that the absolute phase of a ψ function is without physical significance, we see that it is not necessary to assume that the result of applying any permutation of like particles P_r to a legitimate ψ function is to leave it unchanged. On the contrary it *suffices* to assume that

$$P_r\psi = e^{i\eta_r}\psi. \qquad \eta_r = \text{real constant} \qquad (42\cdot1)$$

It is *not* sufficient to require that instantaneously $|P_r\psi|^2 = |\psi|^2$, thus allowing η_r to be a function of the positional coordinates, for ψ determines probability amplitudes in momentum space and other legitimate, generalized, coordinate spaces whose absolute values must also be unaffected by P_r. Hence it is *a priori* reasonable to restrict the quantity η_r in the above equation to constant values. In fact Witmer and Vinti[1] have shown that, unless it is constant, one can derive from a legitimate ψ function others which are not legitimate by the mere process of permutation and linear combination. We therefore assume that Eq. (42·1) is necessary as well as sufficient.

The condition (42·1) implies that ψ is a simultaneous eigenfunction of every permutation operator involving only like particles. In the case of the elementary permutation operators P_{ij} we know that the only eigenvalues are ± 1 (*cf.* Sec. 40d). Moreover, it is not possible for $e^{i\eta}$ to have the value $+1$ for one interchange and -1 for another involving the same kind of particles. To prove this[1] let us assume that $P_{12}\psi = \psi$;

[1] E. E. Witmer and J. P. Vinti, *Phys. Rev.* **47**, 538 (1935).

$P_{13}\psi = -\psi$. By inspection we see that P_{23} is equivalent to both of the products $P_{13}P_{12}P_{13}$ and $P_{12}P_{13}P_{12}$. According to the first formula, ψ would be symmetric with respect to P_{23}, while the second formula would make it antisymmetric. Hence the assumptions lead to a contradiction. We conclude that *every complete physically admissible wave function must be either symmetric* ($e^{i\eta} = +1$), *or antisymmetric* ($e^{i\eta} = -1$), *with respect to all interchange permutations* P_{ij} *involving a single type of elementary particle.* In Sec. 63b we shall see that *the experimental facts require the choice of wave functions which are antisymmetric with respect to interchanges of electrons and of protons.*[1] *This limitation on complete physically admissible ψ functions is the wave-mechanical form of the Pauli exclusion principle.*

If it were not for this exclusion principle, the proof given in Sec. 15 that $\Psi(x,t)$ is in principle an experimentally measurable property of a pure-state assemblage would fail, owing to our inability to distinguish between configurations which differ from one another only by a permutation of like particles. In view of the exclusion principle, however, we know that $|\Psi|^2$ has the same value at all such points of configuration space and can determine its value at any individual point from the measured value of the sum over the set of points obtained from the original one by permutations. Hence $|\Psi|^2$, $d|\Psi|^2/dt$ and ultimately Ψ itself are experimentally determinable.

Let us now return to the question of symmetric and nonsymmetric dynamical variables. Consider an operator α which is given as an explicit function of the coordinates x_i and momenta p_i of the individual particles:

$$\alpha = f(x_1,p_1; x_2,p_2; \cdots).$$

Let the function f have such a form that

$$P_r f(\xi_1,\eta_1; \xi_2,\eta_2; \cdots) = f(\xi_1,\eta_1; \xi_2,\eta_2; \cdots),$$

where ξ_i, η_i are numbers substituted for the operators x_i, p_i. In this case we say that α is *symmetrical.* Every operator which satisfies this requirement must commute with all such permutations P_r. It follows that, if ψ_n is an eigenfunction of α with the eigenvalue a_n, $P_r\psi_n$ is another eigenfunction with the same eigenvalue.

Every P_r can be regarded as the product of a number of interchange operators P_{ij}. The resolution into these elementary factors is not unique, but the number of factors is always even, or always odd. We accordingly define a permutation as even or odd, according as it is composed of an even or odd number of interchanges. Let the index p_r be defined as zero, if P_r is even, and, as unity, if P_r is odd. Denoting the number of

[1] *Cf.* W. HEISENBERG, *Zeits. f. Physik* **38**, 411 (1926).

like particles (electrons or protons) as before by f, we consider the operator[1]

$$\mathcal{G} = \frac{1}{f!}\sum_{\tau}(-1)^{p_\tau}P_\tau. \tag{42·2}$$

The sum is to be extended over the $f!$ distinct permutations of the f particles.

If we multiply any two permutations together we get a third permutation—this is the fundamental group property. Moreover, if $P_\lambda = P_kP_\tau$, $P_{\lambda'} = P_kP_{\tau'}$, and $P_\tau \neq P_{\tau'}$, it follows that $P_\lambda \neq P_{\lambda'}$. Similarly $P_\tau P_k \neq P_{\tau'}P_k$. Therefore, if we multiply each of the $f!$ permutations in turn by any one, either as a prefactor or a postfactor, the products comprise all $f!$ permutations. If P_k is an even permutation, P_kP_τ is even or odd according as P_τ is even or odd. If P_k is odd, P_kP_τ is even when P_τ is odd and *vice versa*. If we multiply $\mathcal{G}$ by any permutation P_k, each term is converted into one of the other terms with its sign changed if P_k is odd, but with its sign unchanged if P_k is even. Hence

$$P_k\mathcal{G} = \mathcal{G}P_k = (-1)^{p_k}\mathcal{G}.$$

It follows at once that $\mathcal{G}\psi$ either vanishes or is an eigenfunction of every permutation operator P_k with the eigenvalue $+1$ if P_k is even and the eigenvalue -1 if P_k is odd. *If it does not vanish, it is antisymmetric with respect to every one of the interchange permutations P_{ij}.*

If the function ϕ is antisymmetric with respect to all the interchanges P_{ij}, it is unchanged by the application of any even permutation and simply reversed in sign by any odd permutation. Consequently

$$\mathcal{G}\phi = \frac{1}{f!}\sum_{\tau=1}^{\tau=f!}\phi = \phi. \tag{42·3}$$

Thus every antisymmetric function is an eigenfunction of $\mathcal{G}$ with the eigenvalue $+1$. It follows that

$$\mathcal{G}(\mathcal{G}\psi) = \mathcal{G}\psi \tag{42·4}$$

and that

$$\mathcal{G}(\psi - \mathcal{G}\psi) = \mathcal{G}\psi - \mathcal{G}\psi = 0. \tag{42·5}$$

Hence every function of the coordinates of f particles can be resolved into the sum of two terms, viz., $\mathcal{G}\psi$ and $(\psi - \mathcal{G}\psi)$, one of which is antisymmetric and an eigenfunction of $\mathcal{G}$ with the eigenvalue $+1$, while the other is an eigenfunction of $\mathcal{G}$ with the eigenvalue 0.

There is no linear combination of permutations of ψ which is antisymmetric and linearly independent of $\mathcal{G}\psi$. To prove this we assume

---

[1] Except for the normalizing factor, $\mathcal{G}$ is the same as the "antisymmetrizer" of Condon and Shortley, *T.A.S.*, Chap. VI, Sec. 3⁶.

that the function $g\psi \equiv \sum_r c_r P_r \psi$ is antisymmetric. Then $\mathcal{G} g\psi = g\psi$.

Hence $\qquad \mathcal{G} g\psi = \sum_r c_r \mathcal{G} P_r \psi = \left(\sum_r c_r (-1)^{p_r} \right) \mathcal{G} \psi = g\psi$.

Thus $g\psi$ is a multiple of $\mathcal{G}\psi$. *We conclude that $\mathcal{G}\psi$ and its multiples are the only antisymmetric functions that can be formed from the linear manifold of the permutations of ψ.*

If α is any dynamical variable which commutes with all the permutation operators, it commutes with $\mathcal{G}$. Consequently there exists a complete set of simultaneous eigenfunctions of α and $\mathcal{G}$. Every antisymmetric function ϕ, being an eigenfunction of $\mathcal{G}$, can be expanded in terms of antisymmetric eigenfunctions of α. Thus the experimental resolution of an arbitrary, physically admissible, pure-state assemblage into subassemblages whose wave functions are eigenfunctions of α involves no violation of the Pauli exclusion principle. *We conclude that every dynamical variable which commutes with every P_r can be reckoned as symmetric in the coordinates of like particles and as measurable in principle.* On the other hand, it is evident that dynamical variables which do not commute with every permutation P_r cannot commute with $\mathcal{G}$ and cannot be measured without violating the Pauli principle. The symmetric dynamical variables thus defined form the most restricted class of operators concerning which we can say that there is no *a priori* reason why they should not be measurable. Hence we shall refer to them as *observables*, using the latter term in a narrower sense than that of Dirac.

If the observables $\alpha_1, \alpha_2, \alpha_3, \cdots$ form a normal commuting set of type 1 operators,[1] it will be possible to expand any physically admissible wave function into simultaneous eigenfunctions of the set, each of which satisfies the Pauli principle. Such a set of type 1 observables will be called *complete* when any two *antisymmetric* simultaneous eigenfunctions with the same eigenvalues are linearly dependent (*cf.* Sec. 37*d*, p. 287). In place of the requirement for type 1 operators introduced on pp. 249 and 254 (Sec. 36), we now postulate that *every type 1 observable can be united with one or more additional observables of the same variety to form a complete normal commuting set.* The formulation of a corresponding postulate for observables which are not of type 1 can be left to the reader.

As $\mathcal{G}$ commutes with every allowable Hamiltonian function H, it commutes with the time-displacement operator $e^{-\frac{2\pi i t H}{h}}$ (*cf.* Sec. 32*b*, p. 201). If $\Psi(x,t_0)$ is antisymmetric, it follows that $\Psi(x,t_0 + t)$, *i.e.*, $e^{-\frac{2\pi i t H}{h}} \Psi(x,t_0)$, is also antisymmetric. Consequently a wave function which satisfies the Pauli principle at any instant t_0 must satisfy it at every other instant.

Some remarks concerning the observation of the configurations of macroscopic bodies are apropos at this point. When one notes the posi-

---

[1] From the Dirac point of view the restrictions to operators of type 1 is unnecessary.

tion of a chair he is clearly observing a function of the coordinates of the particles that compose it, which is symmetrical with respect to all permutations of like particles. One takes the continuity in time of the image of a chair on his retina as proof that the final image and the initial image refer to the *same* object. The existence of moving pictures shows, however, that the apparent continuity of a succession of images is not proof of their correlation with a single object. If there were many identical chairs in a room, an instantaneous permutation of the coordinates of these chairs would be unobservable. And, even granting the uniqueness of any particular macroscopic object, such as a chair, we can have no evidence that the elementary particles which compose it at the time t' are the same as at some other time t''. Thus the concept of identity rests on continuity of experience and ceases to have meaning when experience is essentially discontinuous. In our experiments with atomic systems, however, discontinuity of experience becomes the rule rather than the exception. In fact, it can be shown that continuity of experience is impossible with reference to observations made on the electrons in the low energy states of atoms and that the concept of the identity of the electrons becomes operationally meaningless for this class of experiments.

In order to obtain an approximately continuous set of atomic-configuration observations it would be necessary to make positional observations without disturbing the momenta by amounts which are appreciable in comparison with their observed values. A measurement of configuration to be used for the identification of the electrons must involve an uncertainty Δq in the position of each one which is small compared with the average distance between electrons and much smaller than the effective atomic radius, which we denote by a.

Let $\bar{T}$ denote the mean kinetic energy of a helium atom. Let $\vec{p_1}$ and $\vec{p_2}$ be the momenta of the electrons and let μ be their mass. Then

$$\bar{T} = \frac{\overline{p_1^2}}{2\mu} + \frac{\overline{p_2^2}}{2\mu}.$$

It follows from the virial theorem (*cf.* p. 506) that the average kinetic energy is the negative of the total energy. Hence the mean square of each component of linear momentum is $\overline{p_q^2} = -\frac{\mu E}{6}$. Let Δp be the root mean square of the momentum perturbation due to the measurement of the conjugate positional coordinate. It should be small in comparison with $-(\mu E/6)^{1/2}$ if successive observations are to reveal a continuous orbit. The potential energy V is $-2e^2\left(\frac{1}{r_1}+\frac{1}{r_2}\right) + \frac{e^2}{r_{12}}$. Its mean value must be of the order of $-\frac{3e^2}{a}$. By the virial theorem, $\bar{V} = 2E$. Thus the observation of a continuous orbit implies that

$$\Delta p \Delta q << a\left(-\frac{\mu E}{6}\right)^{1/2} \cong \left(-\frac{3e^2}{2E}\right)\left(-\frac{\mu E}{6}\right)^{1/2},$$

whereas the Heisenberg uncertainty principle requires that $\Delta p \Delta q \geqslant h/4\pi$. Putting

these two requirements together we see that an experimental continuous orbit can exist only if $-E$ $(<<)^2$ $3Rh$, where R is the Rydberg constant. The actual value of $-E$, however, is of the order of $6Rh$.

42c. A Classification of Observations.—It is normally possible to distinguish between a primary and a secondary observing mechanism. The function of the latter is to detect the effect of the interaction between the observed system and the primary mechanism. If the dynamical variable under observation is conserved during the period of interaction with the primary observing mechanism, this period can be as long as desired. If it is not so conserved, the time of the primary interaction must ordinarily be brief. This is true, for example, in observations of position and momentum by collision. In measurements of momentum by the basic method of Chap. II, Sec. 15, the primary observing mechanism must neutralize any force field acting on the particle under observation, and the momentum is thus conserved during the long time which must elapse between the measurements of position.

Observations can be divided into two classes according as the result is made to hinge on the effect of the primary interaction on the system observed, or on the primary observing mechanism. A spectroscopic observation, in which the primary observing mechanism consists of a slit, a pair of lenses, and a prism, belongs to the first class. We are interested here in the effect of the primary interaction on the photons to be measured, and not, for example, in the momentum which the prism receives from the light which passes through it. Similarly, in the case of the Stern-Gerlach experiment, where we have to do with a stream of atoms interacting with a magnet producing an inhomogeneous field, it is the effect of the interaction on the atoms observed which is important. On the other hand, when we measure the position of an object with an optical device, or when we measure velocity in the line of sight by the Doppler effect, we concentrate our attention on the effect of the interaction on the radiation which constitutes the primary observing mechanism. It is evident that whenever the observed system and the primary observing mechanism have very unequal masses the result of the observation must hinge on the behavior of the lighter of the two.

In case the masses of the system observed and the primary observing mechanism are not very different a complete quantum-mechanical theory of the experiment must treat them as different parts of a single system during the period of interaction, although they must be well separated in the beginning and at the end. If the two masses are quite different, however, we can usually describe the interaction by means of a perturbing term H_1 in the Hamiltonian of the lighter pair of interacting systems. This amounts to dealing with the heavier system on a classical basis.

A particularly important variety of observation for quantum mechanics is that in which the primary observing mechanism is very massive in comparison with the system observed and in which the observable to be measured is conserved during the period of the primary interaction. We shall designate this variety of observation as of type A. Measurements of type A can be repeated because the property to be measured is not destroyed by the interaction.

Let the Hamiltonian for the period of interaction in a type A observation on α be $H = H_0 + H_1$. Since α is conserved, it follows that

$$H\alpha - \alpha H = 0.$$

Let Ψ denote the wave function for the assemblage of systems under observation, or for an hypothetical assemblage prepared like the system to be measured, if there is only one. This function experiences a transformation during the period of primary interaction Δt which we can symbolize by means of the time-displacement operator T. Thus

$$\Psi_{t_0} \to T\Psi_{t_0} = \Psi_{t_0 + \Delta t},$$

where the unitary operator T has the explicit form

$$T = e^{-\frac{2\pi i}{h}\Delta t H} \equiv \sum_{n=0}^{\infty} \left(-\frac{2\pi i \Delta t}{h}\right)^n \frac{H^n}{n!}$$

(*cf.* Sec. 32*b*, p. 201 and Sec. 42*b*, p. 339) and is thus a function of the perturbed Hamiltonian H. It follows that T like H must commute with α.

The effect of the primary interaction, or of the equivalent operator T, is usually to separate the subassemblages in space, so that a subsequent positional observation by the secondary observing mechanism is sufficient to determine the value of α for any individual system.

The magnetic-deflection method of measuring energy discussed in Sec. 41*d* is an example of a type A measurement in which the magnet, which produces the field, and the diaphragm are the essential parts of the primary observing mechanism. The measurement of e/m with a mass spectrograph, the Stern-Gerlach experiment for the measurement of atomic magnetic moments, and spectrum analyses by prism and grating are other examples of the same nature. In all these cases a massive primary mechanism interacts with a stream of atomic systems (including photons), in such a way as to spread the latter out into a spectrum where a subsequent positional observation completes the experiment. The measurement of the momentum of a free particle by the method of Sec. 15 constitutes a variant of this type of observation in which no primary observing mechanism is needed.

42d. Measurements as Correlations.—In the above mentioned experiments the variable to be measured is constant for the unperturbed

system so that the time at which the observation begins need not be specified. On the other hand, if we have to do with a quantity α which is not an integral of the natural motion, *e.g.*, the x coordinate of a particle whose average velocity-component along the x axis is not zero, we must usually specify the time at which the measurement is made. In other words the observation normally consists in a correlation of two different quantities, say a value of α and a time. Sometimes, however, the measurement can involve a correlation of two different variables, say x and y, neither of which is the time.

As an example consider the process of passing an electron or a stream of electrons through a slit or other aperture in a diaphragm. This is a measurement which approximately defines a point in the orbit of each particle passing through the aperture. The interaction between the systems under observation and the primary observing mechanism divides the initial‘assemblage into two subassemblages with mutually orthogonal wave functions and consisting respectively of those systems which pass through the aperture and those which strike the diaphragm. We can regard the observation as one which fixes the value of an observable ξ whose eigenfunctions are of two types, *viz.*, those which vanish with their gradients over the aperture and those which vanish with their gradients over all other parts of the plane of the aperture. These eigenfunctions can be associated with arbitrary corresponding eigenvalues if desired, but the eigenvalues are of no importance in connection with such "yes" or "no" observations as this. The experiment can be classified as a type A observation since the interaction of the particles under observation and the diaphragm can be represented by a perturbing term in the Hamiltonian of the former, and since the variable measured is in a sense conserved (so that a second diaphragm like the first and placed beside it would not affect the result). It is hardly necessary to remark that as the aperture of the diaphragm shrinks in size, the observation of ξ becomes an arbitrarily exact measurement of the positional coordinates for the subassemblage of atomic systems which pass from one side of the diaphragm to the other.

42e. The Observing Mechanism Not Entirely Classical.—Consider next the more general type of observation in which the mass of the primary observing mechanism B is not large compared with that of the observed system A, and in which both A and B must be treated on a quantum-mechanical basis. Observations of the position, energy, and momentum of atomic systems by electron or photon impact are examples. In measurements of this type the systems A and B are allowed to interact and a subsequent measurement of B is interpreted with the aid of one of the conservation laws as a measurement of A.

In such cases the classical point of view demands the existence of a definite history of events which can be divided into three periods, *viz.*,

the time before interaction, the time after the interaction, and the time during which the interaction takes place. A similar historical sequence obtains in quantum theory provided that the times in which the systems *A* and *B* are prepared and measured do not overlap each other or the time of interaction. In practice, experiments of this kind are usually made in such a manner as to spoil this historical sequence, although an equivalent spatial sequence takes its place. Thus, in measuring atomic energies by electron impact we allow a steady stream of electrons of known energy emerging from a slit to pass through a chamber containing low-density gas and observe the absorption of energy by reading micro-ammeters which give the currents to various electrodes in the observation chamber. We have to do with a state of kinetic equilibrium in which the preparation of the incident stream of electrons *B* and the observation of their energies after impact are both continuous processes over-lapping in time. We distinguish between the electrons which have not yet interacted with the atoms, and those which have interacted, by their positions rather than by the clock. This fact is clearly unessential, however, and, since the experiment *could* be carried out with shutters so as to establish a time sequence, we simplify the discussion by assuming that it is actually done in that manner.

Let us further assume for the present that the preparation of the system *A* to be observed and of the system *B* which constitutes the primary observing mechanism is such as to leave each of them in a pure subjective state *before* the interaction. We then have definite initial wave functions $\Phi(\xi,t)$, $X(\eta,t)$ for the separate systems prior to interactions. Here ξ denotes the totality of the coordinates of *A*, and η the totality of the coordinates of *B*. It follows that during this period of time the combined system is in a pure state with the wave function $\Psi = \Phi(\xi,t)X(\eta,t)$. No other assumption is consistent with our knowledge of the independent systems.

The Hamiltonian function of the united system will be of the form $H = H_A + H_B + H'$, where the last term denotes the mutual energy which gives rise to the interaction. The complete wave function Ψ is transformed in time according to the rule $H\Psi = \dfrac{h}{2\pi i}\dfrac{\partial \Psi}{\partial t}$. So long as the systems are known to be far apart, H' plays a negligible role and the assumption $\Psi = \Phi X$ is permissible. During the period of interaction, however, H' spoils the factorization—*i.e.*, ΦX ceases to be a solution of $H\Psi = -\dfrac{h}{2\pi i}\dfrac{\partial \Psi}{\partial t}$. After the interaction we can only say that Ψ is expressible as a sum of products of the form

$$\Psi = \sum\sum c_{nm}\Phi_n(\xi,t)\, X_m(\eta,t), \tag{42·6}$$

where the Φ's and X's are orthonormal sets of solutions of the Schrödinger equations of A and B, respectively, and the c_{nm}'s are constant.[1]

It follows at once that, *although the united system is left in a pure state, the states of A and B taken separately are mixed.* To establish this fact we compute the mean value of a dynamical variable α which "belongs to A," *i.e.*, which is applicable only to functions of the ξ coordinates. Let the wave function of the united system after the interaction be expressed in the form

$$\Psi = \sum_m C_m U_m(\xi,t) X_m(\eta,t), \tag{42·7}$$

where

$$C_m U_m(\xi,t) = \sum_n c_{nm}\Phi_n(\xi,t), \qquad (U_m, U_m) = 1. \tag{42·8}$$

The functions $U_m(\xi,t)$ defined by (42·8) are not orthogonal in general, like the X's. Then

$$\bar{\alpha} = (\alpha\Psi,\Psi)_{\xi+\eta} = \sum_m |C_m|^2 (\alpha U_m, U_m)_\xi, \tag{42·9}$$

where the first scalar product is extended over the coordinate space of the united system while the scalar products which follow the summation sign are extended over the coordinate space of A alone. *Thus the expectation value of the arbitrary operator α is obtained either by averaging over the pure-state assemblage of united systems, or over a mixed assemblage of observed systems A in which the probability of the wave function U_m is*

$$|C_m|^2 = \left(\sum_n c_{nm}\Phi_n, \sum_n c_{nm}\Phi_n\right)_\xi \quad [cf. \text{ Eq. (41·1)}]. \quad Although \ the \ A \ systems$$

have ceased to interact with the B systems, α is not obtainable by treating the A systems as a pure-state assemblage with any assignable single wave function $\Psi(\xi,t)$.

As Schrödinger has pointed out,[2] the subjective state of the united system as defined by its Ψ function does not give us sufficient information to define a corresponding subjective state for each of its independent parts. This is due to the fact that much of the information regarding A contained in the Ψ of the united system has to do with the correlation of

---

[1] Here we use the notation for purely discrete spectra, replacing the continuous-spectrum eigenfunctions with an approximating set of discrete functions as suggested on p. 163. To prove the validity of Eq. (42·6), we assume that the Φ's and X's form complete normal orthogonal sets of eigenfunctions of H_A and H_B, respectively, with appropriate time factors to make them solutions of the corresponding time dependent Schrödinger equations. The product functions will then form a complete set in the coordinate space of the combined system. Neglecting H' we see that the right-hand member of (42·6) is a solution of the Schrödinger equation for the combined system when the c_{nm}'s are constant.

[2] E. SCHRÖDINGER, *loc. cit.*, Sec. 41d, p. 328.

values of the coordinates of A with those of B and is therefore conditional rather than absolute.

It is nevertheless possible to restore the system A to a pure state by making observations on B without directly disturbing the system A itself. For this purpose we must make what is called a complete predictive observation of B, i.e., a simultaneous predictive observation of a complete set of commuting dynamical variables. Such an observation determines the wave function of B, regarded as an independent system, uniquely except for a physically meaningless phase factor. Furthermore, it insures that at the moment when the observation is completed the wave function of the combined system shall have the form

$$(\Psi_{A+B})_{t_0} = u(\xi)\varphi(\eta). \tag{42·10}$$

It follows that so long as there is no further interaction between the two systems, Ψ must remain factorable into a product like that which described the combined system before the first interaction of A and B. Hence an expansion of Ψ like (42·7), after the observation of B, will have but a single term and A is accordingly left in a pure subjective state.

Let us now consider the circumstances under which observations of a dynamical variable β which belongs to B, made after the interaction of A and B, are equivalent to observations of a dynamical variable α characteristic of A. In such cases the variable β can be measured after the initial interaction without further disturbing the system A in any way. As examples the observations of energy, momentum, and position of elementary particles, or of atomic systems, by impact may be mentioned. The possibility of such measurements depends upon the initial preparation of A and B in a special way which permits us to replace the expansion (42·6) by one of a more special character.

We replace the functions $\Phi_n(\xi,t)$ by a complete orthonormal system of eigenfunctions of α, say $\varphi_{nr}(\xi)$. Here the first index n fixes the eigenvalue of α and the second fixes the eigenvalues of such other dynamical variables as must be added to α in order to make a complete set for the system A. Similarly we replace the functions X_m by a complete orthonormal system of eigenfunctions of β, say $\chi_{ms}(\eta)$. The wave function of the combined system, after interaction, can always be written in the form

$$\Psi(\xi,\eta,t) = \sum_{n,r,m,s} c(n,r,m,s,t)\varphi_{nr}(\xi)\chi_{ms}(\eta). \tag{42·11}$$

Let us now suppose that all the coefficients $c(n,r,m,s,t)$ vanish except those allowed by the assumption of a functional relationship between n and m. This will be true only for some special way of preparing the systems A and B. With this assumption, however, we can discard the '

index m in favor of the index n which determines it. The expansion then becomes

$$\Psi(\xi,\eta,t) = \sum_{n,r,s} c(n,r,s,t)\varphi_{nr}(\xi)\chi_{ns}(\eta).$$ (42·12)

A predictive measurement of β now reduces Ψ to the form

$$\Psi = C\sum_{r,s} c(n,r,s,t,)\varphi_{nr}(\xi)\chi_{ns}(\eta),$$ (42·13)

where C is a normalizing factor. This is an eigenfunction of α as well as of β, and hence the predictive measurement of β is also a predictive measurement of α. It does not leave the system A in a pure subjective state unless α is nondegenerate, but leaves it in a mixture of subjective states all of which are eigenstates of α with a common eigenvalue α_n. The probability that a retrospective measurement of β for a system in the subjective state (42·12) will give the value β_n is equal to the probability that a retrospective measurement of α will give the value α_n, *viz.*, $\sum_{r,s} |c(n,r,s,t)|^2$. Hence a retrospective measurement of β is equivalent to a retrospective measurement of α.

As regards actual schemes for preparing the systems A and B so as to reduce the expansion (42·7) to the form (42·12), we consider only a special case. Measurements of energy and momentum by impact methods are based on the corresponding conservation laws. Let α and β be the energies of A and B, respectively. If α and β have unique values before impact, their sum has a unique value E before and after impact. All terms of (42·11) vanish except those for which $\alpha_n + \beta_m$ is equal to E. In the typical case, A is an atom initially at rest in its normal state and B is an electron shot out of an electron gun with known initial energy. By measuring the energy loss of the electron we are able to determine the possible increments in the energy of the atom.

CHAPTER X

MATRIX THEORY

43. MATRIX ALGEBRA[1]

Matrix algebra constitutes an important aid to the study of problems in quantum mechanics and in this section we shall briefly review the elements of the subject, reserving applications for discussion in Secs. 44 and 45.

Let **A**, or $\|A(m,n)\|$, denote the matrix, or two-dimensional array of numbers[2]

$$\mathbf{A} \equiv \|A(m,n)\| \equiv \begin{Vmatrix} A(1,1) & A(1,2) & . & . & . \\ A(2,1) & A(2,2) & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \end{Vmatrix}.$$

If the number of rows and number of columns are both finite we say that the matrix is *finite*. Otherwise it is said to be *infinite*. Matrices with an equal number of rows and columns are said to be *square*. Matrices having an infinite number of rows and of columns will be referred to as *infinite square matrices*.

If a matrix **A** has p rows and q columns we shall say that it is a $p \times q$ element matrix. Two matrices are said to be *similar* if they have the same number of rows and the same number of columns—the definition applies even if the number of rows, or the number of columns, or both, are infinite. Only similar matrices can be added and subtracted.

To add or subtract two similar matrices **A** and **B** we add or subtract corresponding elements. Thus

$$(A + B)(m,n) = A(m,n) + B(m,n). \tag{43·1}$$

A matrix, all of whose elements are zero, is called a zero matrix and is denoted by **0**. Two matrices **A** and **B** are equal if, and only if, all their corresponding elements are equal; *i.e.*, if $\mathbf{A} - \mathbf{B} = \mathbf{0}$. The commutative

---

[1] *Cf.* M. BÔCHER, *Higher Algebra*, or BORN and JORDAN, *Elementare Quanten-mechanik*. The author is indebted to the latter text for many suggestions followed in the construction of this section.

[2] The *first* of the indices (m,n) indicates the *row*, and the *second* indicates the *column*. The catch word "Roman Catholic" (R.C.) is helpful in remembering this fundamental convention of our notation.

and associative laws of addition in ordinary algebra apply to the addition of similar matrices.

Let $\mathbf{A}$ be a $p \times q$ element matrix and let $\mathbf{B}$ be a $p' \times q'$ element matrix. If $q = p'$ we define the product $\mathbf{AB}$ by the rule

$$(AB)(m,n) = \sum_k A(m,k)B(k,n). \tag{43.2}$$

Evidently $\mathbf{AB}$ will be a $p \times q'$ element matrix. Finite similar square matrices can be multiplied in either order and yield products similar to the factors. Infinite square matrices can be multiplied in the same way if the infinite series obtained from the rule (43.2) are convergent. Our chief interest is in such similar finite or infinite square matrices and the rest of the following general discussion of matrix algebra refers primarily to them.

On the basis of Eq. (43.2) it can readily be shown that the associative law of ordinary algebra applies to the multiplication of matrices, although the commutative law does not. Thus, in general,

$$\mathbf{AB} \neq \mathbf{BA}.$$

The product of a matrix $\mathbf{A}$ and an ordinary number (scalar) c is the matrix whose typical element is

$$(cA)(m,n) = cA(m,n).$$

A diagonal matrix is a square matrix all of whose elements vanish except those on the principal diagonal. Thus, if $\mathbf{D}$ is a diagonal matrix,

$$\mathbf{D} \equiv \|D_n\delta_{mn}\| \equiv \begin{Vmatrix} D_1 & 0 & 0 & \cdot & \cdot \\ 0 & D_2 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & D_3 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{Vmatrix}.$$

If the nonvanishing elements of the diagonal matrix $\mathbf{D}$ are real and arranged in order of ascending algebraic value, the matrix is said to be an *ordered* diagonal matrix.

A unit matrix $\mathbf{I}$ is a diagonal matrix all of whose elements are unity.

$$\mathbf{I} = \|\delta_{mn}\|.$$

It follows at once that

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}.$$

The *inverse*, or *reciprocal*, of a square matrix $\mathbf{A}$ is indicated by the symbol $\mathbf{A}^{-1}$ and is defined, in case it exists, by the relations

$$\mathbf{AA}^{-1} = \mathbf{I}, \qquad \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}. \tag{43.3}$$

All integral powers of **A** can now be defined by the rules

$$\mathbf{A}^p = \mathbf{AAA} \cdots \qquad \text{to } p \text{ factors,}$$
$$\mathbf{A}^{-p} = (\mathbf{A}^{-1})^p,$$
$$\mathbf{A}^0 = \mathbf{I}.$$

With these definitions we can establish the usual exponential rules

$$\mathbf{A}^\mu \mathbf{A}^\nu = \mathbf{A}^{\mu+\nu}. \qquad (\mathbf{A}^\mu)^\nu = \mathbf{A}^{\mu\nu}.$$

Every finite square matrix **A** has a corresponding determinant which we call det **A**. It follows from the rule for the multiplication of determinants that

$$\det \mathbf{AB} = \det \mathbf{A} \cdot \det \mathbf{B}. \tag{43.4}$$

As the determinant of a unit matrix is unity we conclude from Eqs. (43.3) and (43.4) that if **A** is a finite square matrix with a reciprocal **A**⁻¹,

$$\det \mathbf{A} \neq 0.$$

Conversely, if **A** has a nonvanishing determinant, we can solve either of the equations (43.3) for the elements of the reciprocal **A**⁻¹. In either case, one obtains $A^{-1}(n,m) = a(m,n)/\det \mathbf{A}$, where $a(m,n)$ is the cofactor of $A(m,n)$ in the expansion of det **A**. Hence for such matrices **A**⁻¹**A** = **I** implies **AA**⁻¹ = **I** and *vice versa*. There is no similar simple rule for the existence of a reciprocal to an infinite matrix. Moreover, the rule **A**⁻¹**A** = **I** implies **AA**⁻¹ = **I** is not valid in general for infinite matrices.

There are two classes of matrices with symmetry properties of special importance for the quantum theory, *viz.*, the *Hermitian* and *unitary* matrices. In order to introduce these classes we define the *adjoint matrix* to **A** as a matrix obtained from **A** by replacing each element by its conjugate complex quantity and then interchanging rows and columns. We denote the adjoint of **A** by **A**†. Thus

$$A^\dagger(m,n) = A(n,m)^*. \tag{43.5}$$

An immediate consequence of this definition is that

$$(\mathbf{A} + \mathbf{B})^\dagger = \mathbf{A}^\dagger + \mathbf{B}^\dagger. \tag{43.6}$$

Furthermore,

$$(AB)^\dagger(m,n) = (AB)(n,m)^* = \sum_k A(n,k)^* B(k,m)^* = \sum_k B^\dagger(m,k) A^\dagger(k,n),$$

or

$$(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger. \tag{43.7}$$

Definition: A Hermitian, or self-adjoint, matrix is a square matrix which is equal to its adjoint. It conforms to the equation

$$\mathbf{A}^\dagger = \mathbf{A}, \tag{43.8}$$

and every element is equal to the complex conjugate of the element symmetrically situated with respect to the principal diagonal of the matrix.

From (43·6) it follows that the sum of two Hermitian matrices is Hermitian. Equation (43·7) shows that the product of two Hermitian matrices is Hermitian if, and only if, the matrices commute. On the other hand the symmetrized product $\frac{1}{2}(\mathbf{AB} + \mathbf{BA})$ is always Hermitian if $\mathbf{A}$ and $\mathbf{B}$ are Hermitian.

Definition: A matrix $\mathbf{U}$ is said to be unitary if

$$\mathbf{UU}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{I}, \qquad \text{or} \qquad \mathbf{U}^\dagger = \mathbf{U}^{-1}. \tag{43·9}$$

In general the sum of two unitary matrices is not unitary, but we see from Eq. (43·7) that *if $\mathbf{U}$ and $\mathbf{V}$ are unitary, the products $\mathbf{UV}$ and $\mathbf{VU}$ are also unitary.*

Another immediate consequence of Eq. (43·7) is that *if $\mathbf{A}$ is Hermitian and $\mathbf{U}$ is unitary, the matrix $\mathbf{B}$ defined by*

$$\mathbf{B} = \mathbf{U}^{-1}\mathbf{AU} \tag{43·10}$$

is also Hermitian.

A finite unitary matrix has a determinant of unit absolute value. To prove this statement we note that $\det \mathbf{U}^\dagger = (\det \mathbf{U})^*$, since interchanging the rows and columns of a determinant does not affect its value. Hence

$$|\det \mathbf{U}|^2 = (\det \mathbf{U})(\det \mathbf{U})^* = (\det \mathbf{U})(\det \mathbf{U}^\dagger) = \det (\mathbf{UU}^\dagger) = 1$$

If all the elements of a square matrix are zero except those which lie in a set of consecutive non-overlapping squares spread out along the principal diagonal, the matrix is said to have *step form*. Such a matrix is illustrated in Sec. 48a, p. 390, Fig. 20. A step matrix can also be regarded as a diagonal matrix each element of which is a square matrix. Thus we write

$$\mathbf{A} = \begin{Vmatrix} A^{(1)} & 0 & \cdots \\ 0 & A^{(2)} & \cdots \\ \cdot & \cdots & \cdot \end{Vmatrix}$$

denoting each of the diagonal squares by a corresponding symbol $A^{(r)}$.

If $\mathbf{A}$ commutes with an ordered diagonal matrix $\mathbf{B} = \|b_n \delta_{nm}\|$,

$$(AB)(n,m) - (BA)(n,m) = A(n,m)[b_m - b_n] = 0.$$

Thus $A(n,m)$ must vanish unless $b_n = b_m$, and $\mathbf{A}$ is therefore a step matrix, in which each step is correlated with one value of b_n.

The determinant of a step matrix is evidently equal to the product of the determinants of the individual steps.

Two step matrices are said to be *similar* if each step of one has the same number of rows and columns as the corresponding step of the other. It is not difficult to see that the product of two similar step matrices $\mathbf{A}$ and $\mathbf{B}$ is a step matrix similar to $\mathbf{A}$ and $\mathbf{B}$.

If a unitary matrix is of step form, it follows from (43·9) that each step is a unitary matrix. Consequently the determinant of each step of a unitary matrix is of unit absolute value. A diagonal matrix is a limiting case of a step matrix in which each step contains a single element. Hence the diagonal elements of a diagonal unitary matrix are of unit absolute value.

44. MATRICES AND OPERATORS

44a. The Derivation of Matrices from Operators.—The algebra of square matrices obeys the same fundamental laws as the operator algebra of Sec. 37. Hence it is not surprising to find that we can establish a definite correlation between linear operators and corresponding matrices. In fact operators and matrices give two different realizations of a common abstract mathematical framework. In Secs. 22*d,e* (pp. 116 to 120) we indicated the correlation between the class of functions quadratically integrable over a certain domain and the class of all complex vectors of infinitely many components $a_1, a_2, \cdots$ which yield convergent sums $\sum_{i=1}^{\infty}|a_i|^2$. Just as an operator α can be used to transform a function into another function $\alpha\psi$, so a matrix $\|A(k',k)\|$ can be used to transform a vector $(x_1, x_2, \cdots)$ into another vector by the equations

$$y_k = \sum_{k'=1}^{\infty} x_{k'} A(k',k). \qquad k = 1, 2, \cdots$$

To establish the correlation between operators and matrices we introduce a linear operator α which has an adjoint $\alpha^\dagger$ with respect to a linear manifold of functions C and a domain of integration M in coordinate space (*cf.* definition, Sec. 32*d*, p. 203). Let $\varphi_1(x)$, $\varphi_2(x)$, $\cdots$ be a complete orthonormal sequence of functions in C. In conformity with the formal expansion

$$\alpha\varphi_k(x) = \sum_{k'=1}^{\infty} \varphi_{k'}\alpha(k',k), \qquad (44\cdot1)$$

we define an infinite square matrix $\alpha = \|\alpha(k',k)\|$ by the formula

$$\alpha(k',k) = \int_{\infty} \varphi_{k'}{}^*\alpha\varphi_k d\tau = (\alpha\varphi_k, \varphi_{k'}). \qquad (44\cdot2)$$

α is said to be the matrix representation of the operator α with respect to the given system of basic functions $\varphi_1, \varphi_2, \cdots$. Different basic systems of functions yield different matrix representations, but, to avoid circumlocution when but one such system is under consideration, we shall refer to the corresponding representation of α as *the* matrix of α.

The matrix of the operator adjoint to α is readily seen to be the adjoint $\alpha^\dagger$ of the matrix α. Thus

$$\alpha^\dagger(k',k) = \alpha(k,k')^* = (\varphi_k, \alpha\varphi_{k'}) = (\alpha^\dagger\varphi_k, \varphi_{k'}).$$

Hence Hermitian operators are correlated with Hermitian matrices. It follows from the assumed quadratic integrability of $\alpha\varphi_k$ that the series $\sum_{k'} |\alpha(k',k)|^2$ is convergent for every k.[1] As $\alpha^\dagger\varphi_k$ is also quadratically integrable, we infer that $\sum_{k'} |\alpha(k,k')|^2$ is convergent for every k.

Let ψ and $\vec{\xi}$ denote, respectively, a function which belongs to C and the vector formed from its Fourier coefficients $\xi_k = (\psi, \varphi_k)$. The Fourier coefficients of $\alpha\psi$ with respect to the base functions $\varphi_1, \varphi_2, \cdots$ are given by $(\alpha\psi, \varphi_k) = (\psi, \alpha^\dagger\varphi_k)$. In view of the completeness of the sequence of φ's the scalar product of any two functions over the domain M is equal to the scalar product of the corresponding vectors. Thus we obtain

$$(\alpha\psi, \varphi_k) = \sum_n \alpha(k,n)\xi_n. \tag{44.3}$$

It follows from this formula that *any two operators α and α' having the same matrix α can be identified for physical purposes.* For, so long as we stick to the manifold C, $(\alpha\psi - \alpha'\psi, \varphi_k) = 0$ independent of the choice of ψ and of k. This means that $\alpha\psi$ and $\alpha'\psi$ are physically equivalent (*cf.* Sec. 36d, p. 256). Equation (44.3) and the postulated quadratic integrability of $\alpha\psi$ and $\alpha^\dagger\psi$ also imply that the sums

$$\sum_k |(\alpha\psi, \varphi_k)|^2 = \sum_k \left| \sum_n \alpha(k,n)\xi_n \right|^2, \qquad \sum_k |(\alpha^\dagger\psi, \varphi_k)|^2 = \sum_k \left| \sum_n \alpha(n,k)^*\xi_n \right|^2$$

are convergent.

From the uniqueness of the correlation of matrices and operators we are led to infer that matrices can be used to define operators just as operators define matrices. Let us therefore assume that we have given an arbitrary matrix β such that $\sum_k |\beta(k',k)|^2$ and $\sum_k |\beta(k,k')|^2$ are convergent. Let us further suppose that we have given a vector in Hilbert space $\vec{\eta} = (\eta_1, \eta_2, \cdots)$ such that

$$\sum_n |\eta_n|^2, \qquad \sum_k \left| \sum_n \beta(k,n)\eta_n \right|^2, \qquad \sum_k \left| \sum_n \beta(n,k)^*\eta_n \right|^2$$

[1] This implies by the inequality of Schwarz (Sec. 22, p. 116) that $\sum_{k'} \alpha(k',k)\alpha(k',k'')^*$ is convergent for every pair of values of k and k''.

are convergent. We know from the Fischer-Riesz theorem[1] that there exists an essentially unique quadratically integrable function $y(x)$ such that $(y,\varphi_k) = \eta_k$, for all k. There also exist quadratically integrable functions $u(x)$, $v(x)$ such that

$$(u,\varphi_k) = \sum_n \eta_n \beta(k,n); \qquad (v,\varphi_k) = \sum_n \eta_n \beta(n,k)^*.$$

If we identify u and v with βy and $\beta^\dagger y$, respectively, we define operators β and $\beta^\dagger$ which are adjoint with respect to a linear manifold of functions which includes $y(x)$ and the sequence of φ's.

If the φ's are physically admissible wave functions, the diagonal elements of α are by their definition the mean values of α for the states which the associated functions describe.

If the basic functions $\varphi_1, \varphi_2, \cdots$ of a matrix scheme are eigenfunctions of α,

$$\alpha(k',k) = (\alpha\varphi_k,\varphi_{k'}) = a_k \delta_{kk'},$$

where a_k is the eigenvalue of α for φ_k. Thus the matrix of α in this scheme is diagonal. The requirement that the φ's shall be simultaneous eigenfunctions of a complete set of independent commuting dynamical variables, say $\alpha_1, \alpha_2, \cdots$, determines the φ's uniquely except for physically meaningless phase factors—provided, of course, that the operators $\alpha_1, \alpha_2, \cdots$ have discrete spectra. Hence it is frequently convenient to designate a matrix scheme based on such φ's as a scheme which makes the α's diagonal.

Let $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ denote the matrices of two operators α and β based on the same normal orthogonal system of functions. Then

$$(\alpha + \beta)(k',k) = \alpha(k',k) + \beta(k',k). \tag{44.4}$$

If $\alpha\beta\varphi_k$ is quadratically integrable for all values of k, and α has an adjoint with respect to a manifold of functions that includes not only every φ_k but also every $\beta\varphi_k$, then

$$(\alpha\beta)(k',k) = (\alpha\beta\varphi_k,\varphi_{k'}) = (\beta\varphi_k,\alpha^\dagger\varphi_{k'}).$$

From (22·32) we infer that

$$(\alpha\beta)(k',k) = \sum_{k''}(\beta\varphi_k,\varphi_{k''})(\alpha^\dagger\varphi_{k'},\varphi_{k''})^* = \sum_{k''}\alpha(k',k'')\beta(k'',k). \tag{44.5}$$

It follows from the above hypotheses that the sums $\sum_{k'}|(\alpha\beta)(k',k)|^2$, $\sum_{k'}\left|\sum_{k''}\alpha(k',k'')\beta(k'',k)\right|^2$ are convergent. Conversely, if $\alpha\varphi_k$ and $\beta\varphi_k$ are quadratically integrable for every k, and if the sum $\sum_{k'}\left|\sum_{k''}\alpha(k',k'')\beta(k'',k)\right|^2$ is convergent for every k, there must exist an operator γ defined by

$$(\gamma\varphi_k,\varphi_{k'}) = \sum_{k''}\alpha(k',k'')\beta(k'',k).$$

[1] *Cf.* footnote 1, p. 259.

It can be proved by interchanging the order of two summations that γ is always equal to $\alpha\beta$. Equations (44·4) and (44·5) show that the sum of the matrices of two operators is equal to the matrix of the sum of the operators and that the product of the matrices is equal to the matrix of the product of the operators taken in the same order.

It follows from the definition of the reciprocal of a matrix that if the operator α has a unique reciprocal α^{-1}, the matrix of α^{-1}, in any scheme, is equal to $(\alpha)^{-1}$.

Thus we see that for a very wide class of functions of operators there exist corresponding matrix functions such that

$$\|f(\alpha, \beta, \cdot \cdot \cdot)\| = f(\alpha, \beta, \cdot \cdot \cdot) \qquad (44·6)$$

Incidentally, if an operator α is unitary, *i.e.*, if it has an adjoint $\alpha^\dagger$ which is also its reciprocal, the matrix of $\alpha^\dagger$ will be adjoint and reciprocal to α. Thus *unitary operators always yield unitary matrices*.

It has been assumed hitherto that the members of the normal orthogonal basic set of functions φ_1, φ_2, φ_3, $\cdot \cdot \cdot$ are distinguished by single ordinal number indices. There is no reason, however, why they should not be labeled with the multiple indices, say n,l,m, so useful in connection with problems involving degeneracy. The expansion (44·1) then takes the form

$$\alpha\varphi_{nlm} = \sum_{n',l',m'} \varphi_{n'l'm'}\alpha(n',l',m'; n,l,m), \qquad (44·7)$$

and the multiplication rule (44·5) becomes

$$(\alpha\beta)(n'',l'',m''; n,l,m) = \sum_{n',l',m'} \alpha(n'',l'',m''; n',l',m')\beta(n',l',m'; n,l,m).$$

44b. Canonical Matrix Transformations.—Let us now consider a reversible linear transformation from one normal orthogonal basic system of functions φ_1, φ_2, φ_3, $\cdot \cdot \cdot$ to a second normal orthogonal system ψ_1, ψ_2, ψ_3, $\cdot \cdot \cdot \cdot$. (For convenience the subscripts k and n will be used for the φ's and ψ's, respectively.) Such a transformation is defined by the equations

$$\psi_n(x) = \sum_k \varphi_k(x) U(k,n). \qquad n = 1, 2, 3, \cdot \cdot \cdot \qquad (44·8)$$

The matrix $\mathbf{U} = \|U(k,n)\|$ is called the matrix of the transformation.

Since the φ's and the ψ's both form normal orthogonal systems,

$$(\psi_{n'},\psi_{n''}) = \sum_k U(k,n')U(k,n'')^* = \sum_k U^\dagger(n'',k)U(k,n') = \delta_{n'n''},$$

or

$$\mathbf{U}^\dagger\mathbf{U} = \mathbf{I}. \qquad (44·9)$$

Let us write the inverse transformation in the form

$$\varphi_k = \sum_n \psi_n W(n,k). \tag{44.10}$$

Substitution of (44.8) into (44.10) yields

$$\varphi_k = \sum_{k'} \sum_n \varphi_{k'} U(k',n) W(n,k), \tag{44.11}$$

from which it follows that $\mathbf{UW} = \mathbf{I}$. Multiplying each side of Eq. (44.9) on the left by $\mathbf{U}$ on the right by $\mathbf{W}$, we obtain

$$\mathbf{UU^\dagger} = \mathbf{I}. \tag{44.12}$$

Thus the matrix $\mathbf{U}$ defined by (44.8) is *unitary*, $\mathbf{W} = \mathbf{U^\dagger} = \mathbf{U^{-1}}$.

The ordered set of complex numbers $U(1,n)$, $U(2,n)$, $U(3,n)$, $\cdots$ constitutes a complex vector $\vec{U}_n$ in a space of infinitely many dimensions.[1] From (44.8) we learn that it is a vector representation of the function ψ_r based on the φ functions as unit vectors. $\mathbf{U}$ is compounded from the totality of these $\vec{U}_n$ vectors, one for each column, as n runs through all possible values. Equation (44.9) states that the vectors form a normal orthogonal set. Equation (44.12) in turn states that the vectors defined by the rows of the matrix $\mathbf{U}$ also form a normal orthogonal set. These two sets of relations constitute a parallel to the nine relations between the direction cosines in a rotation of the reference axes in ordinary three-dimensional space. Thus the unitary transformation (44.8) is a generalization of the homogeneous orthogonal linear transformation of elementary Euclidean geometry. We call it a "rotation of the axes in function space."

Let ψ_A be an arbitrary wave function with the expansion

$$\psi_A = \sum_k \varphi_k \xi_k = \sum_n \psi_n \eta_n. \tag{44.13}$$

Here ξ_k, η_n are simply the appropriate Fourier coefficients of ψ_A. Introducing the transformation (44.8) and its inverse, we readily derive

$$\xi_k = \sum_n U(k,n)\eta_n; \qquad \eta_n = \sum_k U^{-1}(n,k)\xi_k. \tag{44.14}$$

Let $\vec{\xi}$ and $\vec{n}$ denote the vector representations of ψ_A formed from the φ_k's and the ψ_n's, respectively, and treated as one-column matrices. The transformation equations (44.14) then take the form

$$\vec{\xi} = \mathbf{U}\vec{n}; \qquad \vec{n} = \mathbf{U^{-1}}\vec{\xi}. \tag{44.15}$$

[1] *Cf.* Sec. 22e, p. 119.

The scalar product of any two functions ψ_A, ψ_B which are expansible in terms of the basic set of functions φ_1, φ_2, φ_3, $\cdot$ $\cdot$ $\cdot$ is equal to the scalar product of the corresponding vectors (*cf.* Sec. 22, p. 120). It follows that the scalar product of two such vectors, $\vec{\xi}_A$ and $\vec{\xi}_B$, is independent of the base system and hence invariant of a transformation from one base system to another. This invariancy of the scalar product justifies our characterization of this type of transformation as a generalization of the rotation of axes in elementary Euclidean geometry. It also shows that there is a fundamental harmony between our definition of a unitary matrix and the definition of a unitary transformation on p. 247. In the latter place we defined a unitary transformation as the application to a function of a reversible operator which preserves scalar products. We now find that a homogeneous linear transformation from one set of basic functions to another yields a transformation of vector representations of functions which preserves scalar products provided that the matrix of the transformation is unitary. Thus we refer to the transformation described by Eqs. (44·8), (44·14), and (44·15) as a unitary transformation. But a vector may be regarded as a function of a variable, say n, defined only for integral values of n. Hence Eqs. (44·15) form a parallel to Eqs. (36·2) with $[\mathbf{U}\times]$ playing the part of a unitary operator.

Let us now examine the corresponding formulas for the transformation of the matrix of an operator α. We assume that α has an adjoint $\alpha^\dagger$ with respect to a linear manifold of functions which includes all the φ's and all the ψ's. Let $\boldsymbol{\alpha}^{(\psi)}$ and $\boldsymbol{\alpha}^{(\varphi)}$ denote respectively the matrices of α for the ψ and φ systems of base functions. It follows from the completeness relation (22·32) that

$$\alpha(n',n'')^{(\psi)} = (\alpha\psi_{n''},\psi_{n'}) = \sum_{k'}(\alpha\psi_{n''},\varphi_{k'})(\psi_{n'},\varphi_{k'})^*. \qquad (44\cdot16)$$

But $(\psi_{n'},\varphi_{k'})^*$ is equal to $U^\dagger(n',k')$ or to $U^{-1}(n',k')$. Also

$$(\alpha\psi_{n''},\varphi_{k'}) = (\psi_{n''},\alpha^\dagger\varphi_{k'}) = \sum_{k''}(\psi_{n''},\varphi_{k''})(\alpha^\dagger\varphi_{k'},\varphi_{k''})^*$$

$$= \sum_{k''}U(k'',n'')\alpha(k',k'')^{(\varphi)}.$$

Hence

$$\alpha(n',n'')^{(\psi)} = \sum_{k',k''}U^{-1}(n',k')\alpha(k',k'')^{(\varphi)}U(k'',n''). \qquad (44\cdot17)$$

This equation has the equivalent form[1]

$$\boldsymbol{\alpha}^{(\psi)} = \mathbf{U}^{-1}\boldsymbol{\alpha}^{(\varphi)}\mathbf{U}. \qquad (44\cdot18)$$

[1] Equation (44·18) differs in form from Eq. (77) of Kemble-Hill, *loc. cit.*, footnote 1, p. 290, because of a difference in the convention regarding the numbering of the elements of the transformation matrix.

We call $\alpha^{(\psi)}$ the transform of $\alpha^{(\varphi)}$ by the unitary matrix $\mathbf{U}$. It follows from (44·18) that

$$\alpha^{(\varphi)} = \mathbf{U}\alpha^{(\psi)}\mathbf{U}^{-1}. \tag{44·19}$$

[The reader will note the parallelism of Eqs. (36·5) and (44·19).]

Applying the theorem (43·7) to the right-hand member of Eq. (44·18) we obtain

$$(\alpha^{(\psi)})^\dagger = \mathbf{U}^{-1}(\alpha^{(\varphi)})^\dagger\mathbf{U}. \tag{44·20}$$

Hence the transforms of two adjoint matrices are themselves adjoint. If $\alpha^{(\varphi)}$ is Hermitian it follows from this that $\alpha^{(\psi)}$ is also Hermitian (*cf.* p. 351).

The transformation of the matrix representation of a dynamical variable from one set of basic coordinate functions to another by means of a unitary matrix $\mathbf{U}$ in accordance with Eqs. (44·18) and (44·19) is called a *canonical matrix transformation*. It is really a special case of the canonical operator transformation defined in Sec. 36c, p. 247.

The functional relation between different matrices is not altered by the application of a canonical transformation. Thus

$$\mathbf{U}^{-1}(\alpha + \beta)\mathbf{U} = \mathbf{U}^{-1}\alpha\mathbf{U} + \mathbf{U}^{-1}\beta\mathbf{U},$$
$$\mathbf{U}^{-1}\alpha\beta\mathbf{U} = (\mathbf{U}^{-1}\alpha\mathbf{U})(\mathbf{U}^{-1}\beta\mathbf{U}).$$

Hence if $\mathbf{f}$ denotes any function of the matrices α and β formed by repeated addition and multiplication (including multiplication by an ordinary complex number),[1]

$$\mathbf{U}^{-1}\mathbf{f}(\alpha, \beta, \cdots)\mathbf{U} = \mathbf{f}(\mathbf{U}^{-1}\alpha\mathbf{U}, \mathbf{U}^{-1}\beta\mathbf{U}, \cdots). \tag{44·21}$$

As a corollary on Eqs. (44·20) and (44·21) we conclude that if a matrix α is unitary, its transform $\mathbf{U}^{-1}\alpha\mathbf{U}$ by any other unitary matrix $\mathbf{U}$ is also unitary.

Although in general the matrices of quantum mechanics have an infinite number of rows and columns, we frequently have to do with problems in which finite matrices appear and are subject to unitary transformations. Hence it is of importance to know that *the sum of the diagonal terms of a finite square matrix is invariant with respect to a transformation of the form*

$$\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}.$$

To prove this proposition we have only to write out in detail the expression for the transformed value of the sum, *viz.*,

$$\sum_n B(n,n) = \sum_{n,k,k'} S^{-1}(n,k)A(k,k')S(k',n), \tag{44·22}$$

[1] We can easily extend the validity of (44·21) to more general functions. For example, if α^{-1} exists, $\mathbf{U}^{-1}\alpha^{-1}\mathbf{U} = (\mathbf{U}^{-1}\alpha\mathbf{U})^{-1}$, as is immediately verified; furthermore, if $\alpha^2 = \beta$, *i.e.*, $\alpha = \sqrt{\beta}$, then $\mathbf{U}^{-1}\alpha\mathbf{U}\mathbf{U}^{-1}\alpha\mathbf{U} = \mathbf{U}^{-1}\beta\mathbf{U}$, or $\mathbf{U}^{-1}\alpha\mathbf{U} = \sqrt{\mathbf{U}^{-1}\beta\mathbf{U}}$.

reverse the order of summation and apply the relation

$$\sum_n S(k',n)S^{-1}(n,k) = \delta_{k'k}.$$

The diagonal sum of a finite square matrix **A** is called the *spur* or *trace* of the matrix and is indicated by the symbol Spur **A**.

We are now prepared to consider from the matrix point of view the problem of determining the discrete eigenvalues of a dynamical variable together with the corresponding eigenfunctions.

44c. Matrix Form ·of the Eigenvalue-eigenfunction Problem.—We assume that the linear operator α has an adjoint $\alpha^\dagger$ with respect to a linear manifold A which includes the complete normal orthogonal system of functions φ_1, φ_2, $\cdots$ and seek to find solutions of the eigenvalue-eigenfunction equation

$$\alpha\psi(x) = a\psi(x) \qquad (44\cdot23)$$

which belong to the manifold A. We call such solutions "discrete eigenfunctions" of A. Using the notation of (44·13) we denote the Fourier coefficients of the desired function ψ with respect to φ_k by ξ_k. Then

$$(\alpha\psi,\varphi_k) = a\xi_k = (\psi,\alpha^\dagger\varphi_k).$$

Since the φ's form a complete system we can evaluate the last scalar product by means of the Fourier coefficients of ψ and $\alpha\varphi_k$. Thus

$$a\xi_k = \sum_n \xi_n\alpha^\dagger(n,k)^* = \sum_n\alpha(k,n)\xi_n. \qquad k = 1, 2, 3, \cdots \quad (44\cdot24)$$

This is equivalent to the matrix equation

$$\alpha\vec{\xi} = a\vec{\xi}, \qquad (44\cdot25)$$

which is the matrix form of (44·23). The number a is referred to as an eigenvalue of the matrix α as well as of the operator α.

The components of α can be worked out by quadrature when α is defined and the basic set of φ's has been chosen. Hence α may be considered known. There are an infinite number of Eqs. (44·24) corresponding to the infinite number of unknown components of $\vec{\xi}$. The eigenvalue a is also unknown. It follows from our method of derivation that if α has discrete eigenfunctions there must exist corresponding eigenvalues and vectors $\vec{\xi}$ which satisfy (44·25). The latter are called *eigenvectors* of the matrix α. In the case of every eigenvector $\sum_k |\xi_k|^2$ is convergent.

If ψ is normalized, its vector representation must also be normalized according to the rule $\sum_k |\xi_k|^2 = 1$. This is always possible as (44·25)

determines $\vec{\xi}$ at most to an arbitrary constant factor. Conversely, if we find a solution $\vec{\xi}$ of (44·25) with convergent sums

$$\sum_k |\xi_k|^2, \qquad \sum_k \left|\sum_n \alpha(k,n)\xi_n\right|^2, \qquad \sum_k \left|\sum_n \alpha(n,k)^*\xi_n\right|^2,$$

it is easy to prove (*cf.*, p. 353) that a corresponding function $\psi(x)$ exists which belongs to the adjoint manifold of α, has the Fourier coefficients $(\psi,\varphi_k) = \xi_k$, and has the property of reducing all corresponding Fourier coefficients of the two sides of (44·23) to equality. This does not prove that the two sides of (44·23) are equal at every point, but we can still count the function ψ as an eigenfunction of α for quantum-mechanical purposes. The series $\sum_k \varphi_k \xi_k$ always converges in the mean upon this function and will ordinarily converge upon it at every point.

If the matrix α is Hermitian, its eigenvalues are real. Thus $\alpha\vec{\xi} = a\vec{\xi}$ implies $\vec{\xi}^\dagger \alpha^\dagger = a^* \vec{\xi}^\dagger$. Hence

$$\vec{\xi}^\dagger \alpha \vec{\xi} - \vec{\xi}^\dagger \alpha^\dagger \vec{\xi} = (a - a^*)\vec{\xi}^\dagger \vec{\xi}. \tag{44·26}$$

If α is Hermitian the left-hand side vanishes. Since $\vec{\xi}^\dagger \vec{\xi}$, or $\sum_k |\xi_k|^2$, cannot vanish, it follows that $a = a^*$, or that a is real. If the matrix α is unitary, its eigenvalues have unit absolute value, for in this case $\alpha\vec{\xi} = a\vec{\xi}$ implies $\vec{\xi}^\dagger \alpha^{-1} = a^* \vec{\xi}^\dagger$. Hence

$$0 = \vec{\xi}^\dagger \alpha^{-1} \alpha \vec{\xi} - \vec{\xi}^\dagger \vec{\xi} = (aa^* - 1)\vec{\xi}^\dagger \vec{\xi}. \tag{44·27}$$

In either case eigenvectors belonging to different eigenvalues must be orthogonal. Thus if α is Hermitian and $\vec{\xi}_1$, $\vec{\xi}_2$ are eigenvectors with the eigenvalues a_1, a_2, respectively,

$$0 = \vec{\xi}_2^\dagger \alpha \vec{\xi}_1 - \vec{\xi}_2^\dagger \alpha^\dagger \vec{\xi}_1 = (a_1 - a_2^*)\vec{\xi}_2^\dagger \vec{\xi}_1 = (a_1 - a_2^*)(\vec{\xi}_1,\vec{\xi}_2). \tag{44·28}$$

If α is unitary

$$0 = \vec{\xi}_2^\dagger \alpha^{-1} \alpha \vec{\xi}_1 - \vec{\xi}_2^\dagger \vec{\xi}_1 = (a_2^* a_1 - 1)\vec{\xi}_2^\dagger \vec{\xi}_1. \tag{44·29}$$

It is always possible to form a set of n mutually orthogonal and normalized vectors $\vec{\xi}_1, \vec{\xi}_2, \cdots, \vec{\xi}_n$ from any set of n linearly independent vectors $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n$ by appropriate linear combination.[1] Since

---

[1] *Cf.* COURANT-HILBERT, *M.M.P.*, 2d ed., Chap. II, p. 41, for procedure.

any linear combination of solutions of (44·25) is a solution, we can always form n orthogonal eigenvectors from any n linearly independent eigenvectors with a common eigenvalue.

Usually the actual solution of the infinite set of homogeneous linear equations (44·24) is very difficult and we have to be content with approximate solutions obtained by perturbation methods (*cf.* Chap. XI). Since the problem is an obvious extrapolation of the simpler one of finding the eigenvalues of a finite square matrix, it is best to examine this latter problem first, especially as the solution of such a simplified case is the usual first step in dealing with the infinite case. We accordingly seek eigenvectors of the $g \times g$ element matrix $\mathbf{A} = \|A_{mn}\|$. Let $\vec{\mathbf{x}} = \|x_k\|$ denote such a vector. By (44·24) its components must yield a nontrivial (*i.e.*, nonvanishing) solution of the set of g equations

$$\sum_n (A_{mn} - a\delta_{mn})x_n = 0. \qquad m = 1, 2, \cdots, g \qquad (44\cdot30)$$

Such a solution exists only if the determinant of the coefficients vanishes. *i.e.*, if a is a root of the so-called "secular" equation

$$\det (\mathbf{A} - a\mathbf{I}) = \begin{vmatrix} A_{11} - a & A_{12} \; \ldots \ldots \ldots A_{1g} \\ A_{21} & A_{22} - a \ldots \ldots \\ . & . \\ . & . \\ . & . \\ . & . \\ . & . \\ A_{g1} \ldots \ldots \ldots \ldots A_{gg} - a \end{vmatrix} = 0. \quad (44\cdot31)$$

This equation is of degree g in the unknown a and has accordingly g roots. If we insert for a in Eqs. (44·30) a root of (44·31) of multiplicity p, the matrix of the coefficients in the former set of equations has its rank reduced to $g - p$. There are then p linearly independent solutions which are readily derived by methods described in texts on algebra. From them we can form p orthogonal eigenvectors in an infinite number of ways, if $p > 1$.

In the more difficult case of Eq. (44·25) where the matrix $\boldsymbol{\alpha}$ has an infinite number of rows and columns and the eigenvectors have an infinite number of elements, the solutions cannot be found so easily, nor can we be sure *a priori* that they even exist. Intuition would suggest, however, the possibility of dealing with (44·25) by means of a method of successive approximations in which at each step we have to do with finite matrices obtained from $\boldsymbol{\alpha}$ and $\vec{\xi}$ by arbitrarily limiting the number of rows and columns considered. For the present we reserve further discussion of methods for solving (44·25).

If we find a normal orthogonal set of eigenvectors of α and designate them by $\vec{\xi}^{(1)}, \vec{\xi}^{(2)}, \vec{\xi}^{(3)}, \cdots$, preferably so as to bring the vectors with a common eigenvalue together and to arrange the eigenvalues in the order of their algebraic magnitude (assuming the eigenvalues to be real), we have the material for the construction of a matrix

$$\mathbf{U} = \begin{Vmatrix} \xi_1^{(1)} & \xi_1^{(2)} & \ldots \ldots \\ \xi_2^{(1)} & \xi_2^{(2)} & \ldots \ldots \\ \ldots & \ldots \ldots \ldots \end{Vmatrix} \equiv \begin{Vmatrix} U(1,1) & U(1,2) \ldots \\ U(2,1) & U(2,2) \ldots \\ \ldots \ldots \ldots \end{Vmatrix} ,$$

such that $\mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$. If a_n denotes the eigenvalue which goes with the eigenvector $\vec{\xi}^{(n)}$ and the eigenfunction ψ_n, Eq. (44·24) takes the form

$$\sum_m \alpha(k,m)U(m,n) = a_n U(k,n). \tag{44·32}$$

Denoting the diagonal matrix $\|a_n\delta_{nm}\|$ by $\mathbf{A}$, we can gather together the different equations (44·24) for the different eigenvectors into the single equation

$$\alpha\mathbf{U} = \mathbf{U}\mathbf{A}. \tag{44·33}$$

Let us now introduce the assumption that the operator α has a complete orthonormal system of discrete eigenfunctions. The φ functions will then be expansible in terms of the discrete eigenfunctions, ψ_n, of α and the transformation

$$\psi_n = \sum_k \varphi_k U(k,n)$$

will be reversible. Under these circumstances the matrix $\mathbf{U}$ has a reciprocal, viz., the matrix of the inverse transformation, and is unitary. Multiplying each side of (44·33) by $\mathbf{U}^{-1}$, we obtain

$$\mathbf{U}^{-1}\alpha\mathbf{U} = \mathbf{A}. \tag{44·34}$$

If a unitary matrix $\mathbf{U}$ can be found which satisfies Eq. (44·34), it must contain all the eigenvectors of (44·25) and by (44·19) *it reduces α to a diagonal form which contains all the eigenvalues.* The eigenvalue problem of matrix theory is usually stated as that of finding such a matrix. If, however, the operator α has a partially continuous spectrum, its discrete eigenfunctions will not form a complete set and no proper unitary matrix $\mathbf{U}$ which satisfies (44·34) exists. In cases of this kind Eqs. (44·24) and (44·25) are still useful and have solutions which describe the incomplete discrete spectrum of α.

A transformation such as (44·34) which reduces a matrix α to diagonal form is frequently referred to as a *principal-axis transformation* since it converts the quadratic form $\xi^\dagger\alpha\xi = \sum_{ij} \alpha_{ij}\xi_i^*\xi_j$ into a sum of squares

and so parallels the reduction of the equation for a quadric surface in three dimensions to its principal axes.

The matrix U of Eqs. (44·33) and (44·34) is not uniquely defined, for, as previously stated, a unitary transformation of the eigenvectors belonging to any given eigenvalue among themselves yields a new set of eigenvectors as good as the first. Let U_1 and U_2 denote any two solutions of Eq. (44·34). Then $U_2 = U_1 V$ where V is a unitary matrix. By hypothesis

$$U_1^{-1} \alpha U_1 = A_1, \qquad U_2^{-1} \alpha U_2 = A_2,$$

where A_1 and A_2 are diagonal. Then

$$V^{-1} A_1 V = A_2.$$

Multiplying through by the prefactor V and writing out the expression for the typical element of each side of the resulting equation, we obtain

$$A_1(k,k) V(k,k') = V(k,k') A_2(k',k').$$

It follows that all elements of V must vanish except those which relate rows and columns belonging to a common eigenvalue of α. Conversely any unitary matrix with this property transforms one solution of (44·34) into another. If we apply the transformation

$$\psi_n^{(2)} = \sum_k \psi_k^{(1)} V(k,n)$$

to the functions $\psi_k^{(1)} = \sum_k \varphi_{k'} U_1(k',k)$ we shall simply form new eigenfunctions of α by taking linear combinations of the old belonging to a common eigenvalue. If the numbering of the matrix U is such that A_1 is an ordered diagonal matrix (*i.e.*, one with all equal values of the diagonal elements grouped together), the matrix V will have a step form such that if we think of V as laid on top of A_1 the steps of V will be squares built about corresponding sets of equal diagonal elements of A_1.

Each step of such a unitary matrix V will have a determinant of unit absolute value (*cf.* p. 351). If the eigenvalues of α are nondegenerate, V will be diagonal and its nonvanishing elements will be of unit absolute value.

*44d. Matrices with Continuous Elements.—The matrix formulation of the eigenvalue-eigenfunction problem is so useful that it is frequently desirable to extend this type of formalism to problems involving continuous spectra. One way to do this is to follow a procedure used in Secs. 30 and 32*j*, introducing in the beginning an arbitrary bounding surface, or box, on which the wave function is required to vanish, carrying the problem through to the point of deriving the results to be compared with experiment and then in the final formulas allowing the surface of the box to move out to infinity.[1]

Another method in which Dirac has led the way is to extend the concept of a matrix to include functions of two independent variables,

[1] The problem has been admirably discussed by J. Frenkel, *Wave Mechanics, Advanced General Theory*, Secs. 10 and 14, Oxford, 1934.

or sets of variables, which take on continuous as well as discontinuous ranges of values. This procedure involves difficulties of the same kind as those met in trying to expand wave functions in terms of eigenfunctions of multiplication operators (*cf.* Secs. 36*g,h*). These difficulties are overcome as before by the introduction of pathological functions which disregard the convergence troubles which arise when we attempt to reverse the order of certain limiting processes. A formalism is thus set up which permits one readily to carry through the early stages of a calculation without merging the continuous spectrum with the discrete spectrum. Then, at the end of the calculation in the final formulas to be compared with experiment, the proper order of the limiting processes is restored, *i.e.*, the formulas are reinterpreted by using eigendifferentials instead of wave functions and taking the limits as the range of the eigendifferentials approaches zero.

In order to introduce this Dirac method we revert to his notation (*cf.* Sec. 36*g*) using the symbol $(\alpha'|\beta')$ to indicate the probability amplitude in the space α_1', α_2', $\cdots$ for a state in which the variables β_1, β_2, $\cdots$ are known to have the eigenvalues β_1', β_2', $\cdots$. Such a probability amplitude can be regarded as a generalized matrix with rows labeled by the eigenvalues of the α's and columns labeled by the eigenvalues of the β's. In multiplying such matrices we must resort to integration in place of summation to cover the continuous ranges of the independent variables. Thus we define the product of $(\alpha'|\beta')$ into $(\beta'|x')$ as the generalized matrix

$$(\alpha'|x') = \overline{\sum}_{\beta'}(\alpha'|\beta')(\beta'|x').$$

Thus all scalar products can be interpreted as matrix products and the formulas of Sec. 36*i* for the transformation of probability amplitudes from one set of independent variables to another become matrix-product formulas.

The normalization condition for the probability amplitude $(\alpha'|\beta')$ is given by Eq. (36·9) or, implicitly and combined with the orthogonality condition, by Eqs. (36·11) and (36·12). This orthogonality-normalization condition can be used to show that $(\alpha'|\beta')$ regarded as a matrix is *unitary*.

Let $(\alpha'|\beta')_\eta$ denote the mean value of $(\alpha'|\beta'')$ over a hypercube of side η in β'' space with its center at $\beta'' = \beta'$. Equation (36·11) can be given the form

$$\lim_{\eta \to 0} \overline{\sum}_{\beta''} u(\beta'')R_\eta(\beta'',\beta') = u(\beta'), \qquad (44\cdot35)$$

where

$$R_\eta(\beta'',\beta') = \overline{\sum}_{\alpha'}(\alpha'|\beta'')_\eta(\alpha'|\beta')_\eta{}^*. \qquad (44\cdot36)$$

Thus $R_\eta(\beta'',\beta')$ can be regarded as an approximation to a function

$(\beta''|\beta')$ with the property that

$$\overline{\sum}_{\beta''} u(\beta_1'',\beta_2'', \cdots)(\beta_1'',\beta_2'', \cdots |\beta_1',\beta_2', \cdots) = u(\beta_1',\beta_2', \cdots).$$

$$(44\cdot37)$$

If all the β's have continuous spectra, this function $(\beta''|\beta')$ is a product of Dirac δ functions, one for each dynamical variable β_k. If all have discrete spectra, it is a product of the form $\delta_{\beta_1'\beta_1''}\delta_{\beta_2'\beta_2''} \cdots$, where $\delta_{\beta_1'\beta_1''}$ is defined as unity when $\beta_1'' = \beta_1'$ and zero otherwise. Following Dirac we ignore the difficulty of interchanging limiting processes and replace (44·36) formally by

$$\overline{\sum}_{\alpha'} (\beta'|\alpha')(\alpha'|\beta'') = (\beta'|\beta'').$$

$$(44\cdot38)$$

Comparing with (36·75) we see that $(\beta'|\beta'')$ is in fact the probability amplitude in β' space when β_1, β_2, $\cdots$ are known to have the values β_1'', β_2'', $\cdots$. Regarded as a matrix the function $(\beta'|\beta'')$ is diagonal. The diagonal elements in the discrete case are unity. Diagonal elements which are not purely discrete are infinite. Nevertheless $(\beta'|\beta'')$ plays the part of a unit matrix, for if multiplied in either order into any other matrix $g(\beta',\beta'')$ it yields $g(\beta',\beta'')$.

$$\overline{\sum}_{\beta'''} (\beta'|\beta''')g(\beta'''|\beta'') = \overline{\sum}_{\beta'''} g(\beta'|\beta''')(\beta'''|\beta'') = g(\beta'|\beta'').$$

The left-hand member of (44·38) is the typical element of the product of $(\beta'|\alpha')$ by its adjoint $(\alpha'|\beta')$. Hence (44·38) states that $(\alpha'|\beta')$ is unitary.

From this point of view (44·8) is seen as a special case of (36·75) in which the functions $(\beta'|\alpha')$, $(\beta'|x')$, $(x'|\alpha')$ are replaced by $\psi_n(x)$, $\varphi_k(x)$, and $U(k,n)$, respectively. Thus the matrix $\mathbf{U}$ of (44·8) plays the part of a probability amplitude.

Let us turn now to the matrix of an operator $\gamma \equiv \gamma\binom{x}{x}$ with respect to a basic set of orthonormal functions $(x'|\beta')$. We assume that γ has an adjoint $\gamma^\dagger$ with respect to x' space and to a class of functions which includes the quadratically integrable functions $(x'|\beta')$ and also the eigendifferentials of those functions which are not quadratically integrable. If $(x'|\beta')$ is not quadratically integrable in x' space, we cannot expect to expand $\gamma \cdot (x'|\beta')$ in terms of simultaneous eigenfunctions of the β's.[1] However, if $(x'|\beta')_\eta$ is the mean value of $(x'|\beta')$ over a hypercube of side η in β' space (cf. Secs. 36g and 36i), $\lim\limits_{\eta \to 0} (x'|\beta')_\eta = (x'|\beta')$ while $\gamma \cdot (x'|\beta')_\eta$ can be so expanded. This fact is formally represented by the equation

$$\gamma \cdot (x'|\beta') = \overline{\sum}_{\beta''}(x'|\beta'')(\beta''|\gamma|\beta').$$

$$(44\cdot39)$$

[1] Here the dot is introduced after the operator γ to prevent confusion between the transform of $(x'|\beta')$ by γ, *i.e.*, $\gamma \cdot (x'|\beta')$ and a function of γ with the arguments x',β'.

Here the expansion coefficients $(\beta''|\gamma|\beta')$ form a matrix similar to the $\|\alpha(k'',k')\|$ of (44·3). The Dirac notation $(\beta''|\gamma|\beta')$ is here substituted for the notation $\gamma(\beta'',\beta')$ because it gives our equations a particularly symmetrical form. Explicitly $(\beta''|\gamma|\beta')$ has the value

$$(\beta''|\gamma|\beta') = \overline{\sum}_{x'}(\beta''|x')\ \gamma \cdot (x'|\beta') = \int_{\infty}(x'|\beta'')^*\ \gamma \cdot (x'|\beta')dx_1'dx_2' \cdots . \tag{44·40}$$

If $\gamma^{\dagger} = \gamma$, the matrix $(\beta''|\gamma|\beta')$ is evidently Hermitian. We can now generalize Eq. (44·19) to

$$(\alpha'|\gamma|\alpha'') = \overline{\sum}_{\beta''}\overline{\sum}_{\beta'}(\alpha'|\beta'')(\beta''|\gamma|\beta')(\beta'|\alpha''), \tag{44·41}$$

if we replace the ψ_n's by simultaneous eigenfunctions of the α's in x' space and the φ_k's by simultaneous eigenfunctions of the β's.

To obtain the generalized equivalent of the eigenvalue-eigenfunction problem of Eq. (44·34) we identify the operator γ with one of the α's, say α_1. In α' space this operator is multiplicative and, by (44·40), $(\alpha'|\alpha_1|\alpha'')$ is $\overline{\sum}_{x'}(\alpha'|x')\alpha_1 \cdot (x'|\alpha'')$ or $\alpha_1''(\alpha'|\alpha'')$. This is the typical diagonal matrix of the mixed discrete-continuous type which must take the place of A in (44·34). The latter equation is now replaced by

$$\overline{\sum}_{\beta''}\overline{\sum}_{\beta'}(\alpha'|\beta'')(\beta''|\alpha_1|\beta')(\beta'|\alpha'') = \alpha_1''(\alpha'|\alpha''). \tag{44·42}$$

Taking the product of each side by $(\beta'''|\alpha')$ and reducing, we obtain

$$\overline{\sum}_{\beta'}(\beta'''|\alpha_1|\beta')(\beta'|\alpha'') = \alpha_1''(\beta'''|\alpha'') \tag{44·43}$$

as the equivalent of (44·33).

The recognition that the totality of the elements of the transformation matrix U of (44·34) constitutes a probability amplitude is perhaps the point of chief physical interest in the Dirac-Jordan transformation theory.

45. THE MATRIX THEORY OF HEISENBERG, BORN, AND JORDAN[1]

45a. Fundamental Postulates.—In this volume the Schrödinger wave equation is treated as fundamental and matrices are introduced as tools for solving problems based on this equation. In the early formulation of the theory by Heisenberg, Born, and Jordan, however, matrices were primary and had to be dealt with independently of any relationship to a wave equation or to wave functions. The H. B. J. theory was based on

[1] W. Heisenberg, *Zeits. f. Physik* **33,** 879 (1925), M. Born, W. Heisenberg, and P. Jordan, *Zeits. f. Physik* **35,** 557 (1925). *Cf.* also Born-Jordan, *E.Q.*

Bohr's correspondence principle which will be discussed in Sec. 46. We here proceed to a brief summary of the fundamental postulates.

a. The "motions" of a mechanical system with a discrete-energy-level spectrum are to be described by the variation in time of matrices for the coordinates and conjugate momenta. Possible choices of coordinates and momenta are based on classical theory. To distinguish between the matrices defined by Eqs. (44·1) and (44·2) and the matrices here postulated when necessary we shall refer to the former as Schrödinger matrices and to the latter as Heisenberg matrices. The latter will be proved to be a special case of the former.

b. The Heisenberg matrices are Hermitian and have quadratically summable rows and columns.

c. Their elements are exponential harmonic functions of the time. Thus, if q_k is a coordinate and p_k the conjugate momentum, the elements of their matrices satisfy the relations

$$q_k(n,m) = a_k(n,m)e^{2\pi i \nu_{nm} t}, \\ p_k(n,m) = b_k(n,m)e^{2\pi i \nu_{nm} t}. \quad (45\cdot1)$$

The frequencies here introduced are related to the possibility energy levels of the system by the Einstein law

$$h\nu_{nm} = E_n - E_m. \quad (45\cdot2)$$

Thus every row and every column is associated with a definite energy level.

d. The Heisenberg matrices of a system of coordinates $q_1, q_2, \cdots, q_f$ and their conjugate momenta $p_1, \cdots p_f$ are subject to the commutation rule

$$[\mathbf{q}_k, \mathbf{p}_l] = \frac{2\pi i}{h}[\mathbf{p}_l \mathbf{q}_k - \mathbf{q}_k \mathbf{p}_l] = \mathbf{I}\delta_{kl}, \quad (45\cdot3)$$

where $\mathbf{I}$ is a unit matrix. It follows from this relation that the matrices must all be infinite.[1] A system of matrices which conforms to (45·3) is said to be canonical.

e. The Heisenberg matrices are subject to equations of motion of the canonical Hamiltonian form

$$\frac{d\mathbf{q}_k}{dt} = \frac{\partial \mathbf{H}}{\partial \mathbf{p}_k}, \qquad \frac{d\mathbf{p}_k}{dt} = -\frac{\partial \mathbf{H}}{\partial \mathbf{q}_k}, \quad (45\cdot4)$$

where the Hamiltonian matrix $\mathbf{H}$ is a function of the q's and p's and the partial derivatives are defined by the rule

$$\frac{\partial f(\mathbf{x}_1, \mathbf{x}_2, \cdots)}{\partial \mathbf{x}_k} = \lim_{a \to 0}\left[\frac{f(\cdots, \mathbf{x}_k + a\mathbf{I}, \cdots) - f(\cdots, \mathbf{x}_k, \cdots)}{a}\right]. \quad (45\cdot5)$$

[1] *Cf.* Born-Jordan, *E.Q.*, p. 90.

A canonical system of matrices which satisfies Eqs. (45·4) is said to form a canonical solution of the equations of motion.

f. The squares of the absolute values of the elements of the Heisenberg electric-moment matrix

$$\vec{D} = \sum_i e_i \vec{r}_i \tag{45·6}$$

are reflected in the intensities of the spectrum lines when atomic systems emit and absorb radiation in a discharge or absorption tube. The relation between $\vec{D}$ and the line intensities will be more fully discussed in Sec. 54.

45b. Correlation of the Heisenberg and Schrödinger Theories.—In order to relate the two forms of quantum theory we note first of all that postulate **b** is satisfied if we take the **p**'s and **q**'s to be Schrödinger matrices for a set of Hermitian operators whose Hermitian domains include class D and choose the basic orthonormal sequence $\varphi_1, \varphi_2, \cdots$ from that class. Furthermore, the commutation rule **d** becomes a corollary on the corresponding operator rule (37·10) if we require that the operators p_k, q_k shall be canonically conjugate and that the transform of every φ_k by any one of these operators shall belong to the Hermitian domain of the others. If this last condition is not satisfied, the products $\mathbf{p}_k\mathbf{q}_k$ and $\mathbf{q}_k\mathbf{p}_k$ may not exist.[1]

If the q's are taken to be the Cartesian coordinates and the p's their momenta, it follows simply from the definition of class D that the first of the two sufficient conditions for (45·3) is satisfied. This is the usual choice. While other choices may be valid, they must be examined with care, since (45·3) does not hold for the Schrödinger matrices of all pairs of canonically conjugate operators, as was at one time supposed.[2]

In order to satisfy postulate **c** with Schrödinger matrices it is necessary that the phases of the base functions shall be harmonic functions of t. It is *sufficient* to choose for the φ's an orthonormal system of simultaneous solutions of the two Schrödinger equations $H\Psi = E\Psi$ and

$H\Psi = -\dfrac{h}{2\pi i}\dfrac{\partial \Psi}{\partial t}.$ In that case

$$\varphi_n = \psi_n e^{-\frac{2\pi i}{h}E_n t} \tag{45·7}$$

---

[1] It follows from a theorem introduced on p. 354 that the quadratic summability of the rows and columns of the product matrices affords an alternative hypothesis necessary and sufficient to make (45·3) a corollary on (37·10).

[2] Thus Born and Jordan have shown (*cf. E.Q.*, p. 91) that the matrix equation (45·3) cannot hold true if the momentum operator has a purely discrete spectrum and a representation is used in which p is diagonal. A case in point is obtained if we identify p with a component of the angular momentum, say $\mathcal{L}_z$, and q with the conjugate azimuthal angle (*cf.* end of Sec. 39a).

and

$$q_k(n,m) = \int_\infty \varphi_n{}^* q_k \varphi_m d\tau = e^{\frac{2\pi i}{h}(E_n - E_m)t} \int \psi_n{}^* q_k \psi_m d\tau. \qquad (45 \cdot 8)$$

Let us designate these as *canonical Schrödinger matrices.*[1]

We turn next to postulate **e** and the equations of motion (45·4). These equations parallel the operator equations (39·26) just as the definition (45·5) parallels (39·18). Hence we can expect that with suitable restrictions the matrix form of the equations of motion will turn out to be a consequence of the operator form if we use canonical Schrödinger matrices. Actually there is no difficulty in a formal derivation of the matrix equations of motion on this basis. However, the matrix equations can also be derived independently of the operator equations by methods similar to those used in setting up the latter.

The first step in deriving Eqs. (45·4) is to note that from Eq. (44·5) we are led to expect that the functional relationship between H, the Cartesian coordinates x_k, and their momenta p_k is paralleled by a formally identical relationship between the matrices $\mathbf{H}$, $\mathbf{x}_k$, $\mathbf{p}_k$. Thus the operator equation

$$H(x_k, p_k) = \frac{1}{2\mu} \sum_{k=1}^{3} p_k{}^2 - e^2 Z \left[\sum_{k=1}^{3} x_k{}^2 \right]^{-\frac{1}{2}}$$

for the internal energy of an hydrogenic atom leads to the corresponding matrix equation

$$\mathbf{H} = \frac{1}{2\mu} \sum_k \mathbf{p}_k{}^2 - e^2 Z \left[\sum_k \mathbf{x}_k{}^2 \right]^{-\frac{1}{2}}. \qquad (45 \cdot 9)$$

The validity of this equation, like the validity of (45·3), rests upon the existence and quadratic summability of the various terms, but we can introduce the required postulate in good conscience for canonical Schrödinger matrices and Cartesian coordinates. Furthermore, it follows from Eqs. (45·5) and (39·18) that $\partial \mathbf{H}/\partial \mathbf{p}_k$ and $\partial \mathbf{H}/\partial \mathbf{q}_k$ are the matrices of $\partial H/\partial p_k$ and $\partial H/\partial q_k$, respectively. Finally, as the basic functions are solutions of the second Schrödinger equation [cf. Eq. (39·25)] the matrices of the operators dq_k/dt, dp_k/dt are respectively equal to the time derivatives of $\mathbf{q}_k$ and $\mathbf{p}_k$. Thus Eqs. (45·4) are established by equating the matrices of the right- and left-hand members of (39·26).

These considerations suffice to prove that canonical Schrödinger matrices of the Cartesian coordinates satisfy all the conditions imposed on the Heisenberg matrices by postulates **b, c, d, e** above. We have

[1] It is necessary to modify the eigenvalue-eigenfunction problem for H in order to eliminate the continuous spectrum (*cf.* Sec. 32*j*) or to include a continuous portion of each matrix.

yet to consider whether the postulates are sufficient to determine the matrices uniquely or not.

45c. Solution of Matrix Equations of Motion for an Ideal Linear Oscillator.—The problem of the ideal linear oscillator was the first to be solved by the matrix method.[1] In this case the Hamiltonian matrix has the form

$$H(x,p) = \frac{p^2}{2\mu} + \frac{kx^2}{2}. \tag{45·10}$$

The equations of motion reduce to

$$\frac{dx}{dt} = \dot{x} = \frac{p}{\mu}, \qquad \dot{p} = -kx. \tag{45·11}$$

Eliminating p and introducing the classical vibration frequency $\nu_c = (2\pi)^{-1}(k/\mu)^{1/2}$ we obtain

$$\ddot{x} \equiv \|(2\pi i\nu_{mn})^2 x(m,n)\| = -4\pi^2\nu_c^2 x.$$

Hence

$$(\nu_{mn}^2 - \nu_c^2)x(m,n) = 0. \tag{45·12}$$

This equation and (45·2) show that $x(m,n)$ must vanish unless

$$E_m - E_n = \pm h\nu_c. \tag{45·13}$$

It follows that the energy levels, or diagonal elements of H, consist of one or more equally spaced series. In order to obtain a solution of our equations we assume (a) that there is just one such series and (b) that the energy levels are nondegenerate. Since the ordering of the rows and columns of the matrices is immaterial so long as it follows a consistent scheme, we can assume without loss of generality that

$$\nu(n+1,n) = \frac{E_{n+1} - E_n}{h} = \nu_c. \tag{45·14}$$

The nonvanishing elements of x will then be arranged in two lines parallel and adjacent to the principal diagonal. Thus

$$x = \begin{Vmatrix} 0 & x(1,2) & 0 & . & . \\ x(1,2)^* & 0 & x(2,3) & . & . \\ 0 & x(2,3)^* & 0 & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \end{Vmatrix}.$$

By (45·11) the matrix p will have the same form, since

$$p(n,m) = 2\pi i\nu_{nm}\mu x(n,m).$$

[1] M. Born, W. Heisenberg, and P. Jordan, *Zeits. f. Physik* **35**, 557 (1925).

It follows at once that all nondiagonal elements of $[\mathbf{x},\mathbf{p}]$ vanish as required by the commutation rule. The diagonal elements of $[\mathbf{x},\mathbf{p}]$ must have the common value unity. Consequently

$$\frac{h}{8\pi^2\mu} = -\sum_k \nu_{nk}|x(n,k)|^2 = \nu_c[|x(n,n+1)|^2 - |x(n,n-1)|^2]. \quad (45\cdot15)$$

Thus the squares of the absolute values of the matrix elements $x(n,n+1)$ form an arithmetical progression with the common difference $h/8\pi^2\mu\nu_c$. These terms are essentially positive so that n must have a minimum value which we can set equal to zero. For this lowest energy level, $(45\cdot15)$ reduces to

$$|x(0,1)|^2 = \frac{h}{8\pi^2\mu\nu_c},$$

and we obtain for the general case

$$|x(n,n+1)|^2 = (n+1)\frac{h}{8\pi^2\mu\nu_c} \qquad n = 0, 1, 2, \cdots \quad (45\cdot16)$$

For the matrix element $x(n,n+1)$ itself we have

$$x(n,n+1) = x(n+1,n)^* = \left[\frac{(n+1)h}{8\pi^2\mu\nu_c}\right]^{\frac{1}{2}} e^{-i(2\pi\nu_c t + \phi_n)}, \quad (45\cdot17)$$

where the ϕ_n's are arbitrary phase constants.

Having evaluated $\mathbf{x}$ we can next determine $\mathbf{p}$ by $(45\cdot11)$ and $\mathbf{H}$ by $(45\cdot10)$. For the latter we find

$$H(n,m) = E_n\delta_{nm} = (n+\tfrac{1}{2})h\nu_c\delta_{nm}. \quad (45\cdot18)$$

The energies obtained in this way are the same as those worked out by the Schrödinger method in Sec. 20. The matrix components are also readily verified by the Schrödinger method. Thus

$$x(n,m) = e^{\frac{2\pi i}{h}(E_n - E_m)t}\int_{-\infty}^{+\infty} \psi_n^* x \psi_m dx.$$

The substitution

$$\psi_n = c_n H_n(\xi)e^{-\frac{\xi^2}{2}}, \qquad \xi = x\sqrt{\frac{k}{h\nu_c}}$$

yields

$$x(n,m) = \frac{h\nu_c}{k}c_m c_n^* \left[\int_{-\infty}^{+\infty} H_n(\xi)\xi H_m(\xi)e^{-\xi^2}d\xi\right]e^{\frac{2\pi i}{h}(E_n - E_m)t}$$

The recurrence formula $(20\cdot8)$ reduces the integral to a form which is readily evaluated with the aid of $(20\cdot10)$. The calculated values then agree with $(45\cdot17)$.

Born and Jordan[1] have proved that the above solution of the matrix problem is unique except for the physically meaningless phase constants ϕ_n.

45d. Reduction of the Fundamental Problem of Matrix Mechanics to a Principal-axis Transformation.—In the matrix mechanics, as in the classical mechanics, a solution of the equations of motion reduces the Hamiltonian to a constant, provided that $\mathbf{H}(\mathbf{q},\mathbf{p})$ does not contain the time explicitly. This is the matrix form of the law of the conservation of energy. It is proved by the equation

$$\frac{d\mathbf{H}}{dt} = \frac{1}{2}\sum_k \left[\frac{\partial\mathbf{H}}{\partial\mathbf{q}_k}\frac{d\mathbf{q}_k}{dt} + \frac{\partial\mathbf{H}}{\partial\mathbf{p}_k}\frac{d\mathbf{p}_k}{dt}\right] + \frac{1}{2}\sum_k \left[\frac{d\mathbf{q}_k}{dt}\frac{\partial\mathbf{H}}{\partial\mathbf{q}_k} + \frac{d\mathbf{p}_k}{dt}\frac{\partial\mathbf{H}}{\partial\mathbf{p}_k}\right]$$

$$= \sum_k \left[\frac{\partial\mathbf{H}}{\partial\mathbf{q}_k}\frac{\partial\mathbf{H}}{\partial\mathbf{p}_k} - \frac{\partial\mathbf{H}}{\partial\mathbf{p}_k}\frac{\partial\mathbf{H}}{\partial\mathbf{q}_k}\right] = 0. \tag{45.19}$$

Equations (45.1) and (45.2) demand that the diagonal matrix

$$\mathbf{E} = \|E_n\delta_{nm}\|$$

shall satisfy the relations

$$\dot{\mathbf{q}}_k = [\mathbf{q}_k,\mathbf{E}]; \qquad \dot{\mathbf{p}}_k = [\mathbf{p}_k,\mathbf{E}]. \tag{45.20}$$

Furthermore (45.20) implies that if $\mathbf{f}$ is a function of the $\mathbf{q}$'s and $\mathbf{p}$'s

$$\frac{d\mathbf{f}}{dt} = [\mathbf{f},\mathbf{E}]. \tag{45.21}$$

The proof will readily be supplied by the reader following the argument used in establishing Eqs. (39.22) and (39.23). We conclude from (45.19) and (45.21) that if $\mathbf{H}(\mathbf{q},\mathbf{p})$ is built up from a solution of the equations of motion, it must commute with the diagonal matrix $\mathbf{E}$. We can assume without loss of generality that $\mathbf{E}$ is an ordered diagonal matrix. Then, by a rule given on p. 351, $\mathbf{H}$ must be a step matrix in which each step is associated with a single energy E_n.

Let us now apply a unitary transformation

$$\mathbf{Q}_k = \mathbf{U}^{-1}\mathbf{q}_k\mathbf{U}, \qquad \mathbf{P}_k = \mathbf{U}^{-1}\mathbf{p}_k\mathbf{U}. \tag{45.22}$$

Such a transformation preserves the functional relationships between matrices [*cf.* Eq. (44.21)] and hence transforms one canonical solution of the equations of motion into another. We choose $\mathbf{U}$ as a step matrix $\mathbf{V}$ similar to $\mathbf{H}$. Then each step $\mathbf{H}^{(n)}$ will be transformed according to the rule

$$\mathbf{H}^{(n)} \rightarrow (\mathbf{V}^{(n)})^{-1}\,\mathbf{H}^{(n)}\,\mathbf{V}^{(n)}. \tag{45.23}$$

It is always possible to convert a finite symmetric square matrix to

[1] BORN and JORDAN, *E.Q.*, p. 23.

diagonal form (*cf.* p. 361) by means of a properly chosen unitary transformation. Hence we can always choose V so that it makes H diagonal, *i.e., we can always find a principal-axis transformation which converts* H *into* E. Thus if any canonical solution of the equations of motion exists there will be one which makes H diagonal.

Conversely, if we can find a canonical set of matrices which make H diagonal we have a solution of the equations of motion. By (39·22) and (39·23)

$$\frac{\partial H}{\partial q_k} = [H, p_k]; \qquad \frac{\partial H}{\partial p_k} = [q_k, H]. \tag{45·24}$$

If $H(q_k, p_k)$ is diagonal and if the q's and p's are given phase factors in accordance with (45·1) and (45·2)

$$\left(\frac{\partial H}{\partial q_k}\right)(n,m) = \frac{2\pi i}{h}(E_m - E_n)p_k(n,m) = -\left(\frac{dp_k}{dt}\right)(n,m).$$

Similarly the second set of Hamiltonian equations is satisfied, as was to be proved.

This suggests that one way to solve the Heisenberg equations independent of the Schrödinger method would be to start with an arbitrary special canonical system of matrices, say $q_k^{(0)}$, $p_k^{(0)}$, form the corresponding Hamiltonian $H^{(0)} = H(q^{(0)}, p^{(0)})$ and seek a canonical transformation which will diagonalize $H^{(0)}$. Such a transformation, if it can be found, will preserve the canonical character of the q, p matrix system and give a solution of the equations of motion. It cannot always be found without making use of the generalized matrices whose rows and columns are not all discrete. However, this procedure does give the starting point of a systematic method of attacking problems in matrix mechanics which cannot be solved by such a simple direct approach as that used in dealing with the ideal linear oscillator. This is the basis of the perturbation theory of the matrix mechanics.

The question now arises whether the solution of the equations of motion obtained in the above manner will be unique or will depend on the choice of the initial canonical system $q_k^{(0)}$, $p_k^{(0)}$. The point has been investigated by Born and Jordan,[1] who found it possible to insure uniqueness by imposing on the $p^{(0)}$'s and $q^{(0)}$'s the condition that $[p_k^{(0)}]^2 + [q_k^{(0)}]^2$ shall be transformable to diagonal form for every value of k by means of a properly chosen canonical transformation. This condition is equivalent to postulating that all matrices shall be derivable from the matrices which solve the problem of a set of harmonic oscillators, since $p^2 + q^2$ is the Hamiltonian matrix of a harmonic oscillator of properly chosen mass and frequency. From the Schrödinger standpoint we should expect to start with the matrices based on a proper complete orthonormal set of base functions and just that is actually effected by this criterion of Born and Jordan.

[1] BORN and JORDAN, *E.Q.*, p. 128.

46. THE BOHR CORRESPONDENCE PRINCIPLE AND ITS RELATIONSHIP TO MATRIX THEORY

46a. The Bohr Postulates.—Historically the matrix mechanics is an outgrowth of the attempt to refine the correspondence principle of the quasi-classical Bohr theory. As this correspondence principle continues to be a useful tool for the heuristic examination of quantum-mechanical problems we pause here to sketch the Bohr theory and its relation to matrix mechanics.

Bohr's work began with an attempt to reconcile the Rutherford nuclear atomic model with the empirical spectrum of hydrogen. His fundamental assumptions were, briefly, as follows:

a. An atomic or molecular system can exist only in certain discrete nonradiating "stationary states" which define corresponding "allowed" energy levels.

b. Transitions, or "jumps," from one energy level to another take place and are accompanied by the emission or absorption of monochromatic radiation of frequency ν according to the rule

$$h\nu = E' - E'', \tag{46·1}$$

where E' and E'' are respectively the upper and lower energy levels in question.

These were the primary postulates and carry over with some reinterpretations to the quantum mechanics of systems with purely discrete energy spectra. The third hypothesis, however, was of an essentially tentative and provisional character and was introduced not so much because Bohr believed it to be true as because at the time no useful alternative hypothesis suggested itself.

c. When the system is in one of its stationary states the motion of the electrons and other particles which make up an atomic system takes place in accordance with the laws of classical mechanics, *i.e.*, classical electrodynamics with radiation forces omitted.

A fourth basic element in the theory, which took final form some time after the appearance of Bohr's first papers, was the quantization rule. This rule, or quantum condition, first formulated by Sommerfeld, is applicable only to multiply-periodic or "conditionally" periodic systems. These are systems whose motions are such that the variation in time of each coordinate and momentum component can be represented by a multiple Fourier's series with one or more independent basic frequencies $\omega_1, \omega_2, \cdots, \omega_r$, where r is not greater than the number of degrees of freedom. Thus if x_k is one of the coordinates, and if we choose the complex form of Fourier's series, the expansion

$$x_k(t) = \sum_{\tau_1 \cdots \tau_r = -\infty}^{+\infty} x_k(\tau_1, \cdots \tau_r) \exp\left[2\pi i(\tau_1\omega_1 + \cdots \tau_r\omega_r)t\right] \tag{46·2}$$

must be possible when the system is multiply periodic. In this expansion the coefficients $x_k(\tau_1, \cdots \tau_r)$ depend on the constants of integration which fix the "orbit" under consideration.

In the case of such systems it is possible to introduce a set of so-called "action variables" $J_1, J_2, \cdots, J_r$, one for each of the independent basic frequencies, and having the following properties. They are constants of integration which fix the energy E and fulfill the conditions[1]

$$\frac{\partial E}{\partial J_i} = \omega_i, \qquad i = 1, 2, \cdots, r \qquad (46 \cdot 3)$$

$$\sum_{i=1}^{r} J_i \omega_i = 2\bar{T}, \qquad (46 \cdot 4)$$

where $\bar{T}$ is the average value of the kinetic energy of the system over a long period of time. Usually it is possible to choose a set of generalized coordinates $q_1, q_2, \cdots, q_f$ such that each q_k is associated with a definite corresponding fundamental frequency ω_k. The integral $\oint p_k dq_k$ extended over a complete cycle of the variable q_k is called the Sommerfeld phase integral (cf. Sec. 21f) for the coordinate q_k. Where these exist, one can identify each action variable J_k with the sum of the phase integrals for all coordinates with the same frequency ω_k. We can now state the fourth postulate in the following form.

d. The stationary states of a multiply-periodic system comprise those states for which the action variables are integral multiples of Planck's constant h. Thus for these allowed states

$$J_k = \sum_i \oint p_i dq_i = n_k h. \qquad n_k = 0, 1, 2, \cdots \qquad (46 \cdot 5)$$

Later on it was discovered that the empirical facts could be fitted more accurately in some cases if one (or more) of the J's was supposed to be an odd multiple of $h/2$. This modification in the theory is evidently unessential in view of the arbitrary nature of **d.**

46b. The Bohr Correspondence Principle and the Heisenberg Matrix Theory.—An important feature of the Bohr theory is the correlation which it permits between the classical basic frequencies $\omega_1, \omega_2, \cdots, \omega_r$ and the radiated frequencies (quantum frequencies) permitted by the rule (46·1). In order to explain this correlation we make use of an r-dimensional space in which each of the action variables is laid out at right angles to the others. In such a space each allowed energy level is represented by a lattice point defined by (46·5). Using primes to denote the quantum numbers n_k for the upper energy level and double primes for the lower energy level we define a "quantum transition"

[1] *Cf.* J. H. VAN VLECK, *Quantum Principles and Line Spectra*, p. 18, Washington, D. C., 1926.

by the quantum numbers n_1', n_2', $\cdots$, n_r' and by the differences

$$n_1' - n_1'' = \tau_1; \qquad n_2' - n_2'' = \tau_2; \qquad \cdots \; ; \qquad n_r' - n_r'' = \tau_r.$$

The frequency of the radiation emitted in such a jump, *viz.*,

$$\nu = \frac{[E(n_1', n_2', \cdots) - E(n_1' - \tau_1, n_2' - \tau_2, \cdots)]}{h},$$

is readily identified with the average value of the combination frequency

$$\omega^{(r)} \equiv \tau_1 \omega_1 + \tau_2 \omega_2 + \cdots + \tau_r \omega_r$$

evaluated along a line joining those points in J space which represent the initial and final energy levels.[1] Thus every type of quantum transition is correlated with a definite pair of terms in the Fourier expansion $(46 \cdot 2)$[2] and it becomes evident that *if the frequencies become independent of the J values—as in the region of high quantum numbers—the quantum frequencies and the classical frequencies will be the same.*

In order to fit his theory with the classical theory of the radiation of macroscopic systems Bohr superimposed on this correlation principle the assumption that *in the region of high quantum numbers the intensities of the spectral lines become asymptotically equal to the intensity of the radiation computed classically from the corresponding Fourier amplitude in* $(46 \cdot 2)$, *whether for the upper energy level or the lower one. Where the classical intensity for a given harmonic component is zero independent of the state for which it is computed (i.e., independent of the J's), the intensity of the radiation emitted in the corresponding type of quantum jump is zero.*

This is Bohr's correspondence principle and the basis for the theoretical treatment of selection rules governing "allowed" and "forbidden" transitions in the Bohr theory.

The classical rate of emission for a system of charges with a multiply-periodic motion can be resolved into a sum of terms, one for each frequency. The expression for the electric vector in the emitted radiation field can also be broken up into a sum of terms giving the contributions of the electric dipole moment, the magnetic dipole moment, the electric quadrupole moment, etc.[3] Usually the dipole-moment term is so large that it swamps the others out. In that case the rate of emission is given by the familiar formula for an harmonic oscillator of electric moment $\vec{d(t)}$, *viz.*,

$$\text{Power radiated} = \frac{2}{3c^3} \left| \frac{d^2}{dt^2} \vec{d(t)} \right|^2,$$

For a more detailed treatment *cf.* J. H. Van Vleck, *loc. cit.*, pp. 23–28.

    [2] *Two* terms in $(46 \cdot 2)$ obtainable one from the other by reversing the signs of all the τ's, are associated with the same absolute frequency. These terms are conjugate complex quantities and can be united to form a single real harmonic term if desired.

    [3] *Cf.* J. Frenkel, *Elektrodynamik*, Vol .I, p. 158.

provided that we identify $\vec{d}$ with the sum of the terms of appropriate frequency in the Fourier expansion of the electric moment $\vec{D} = \sum_{k} e_k \vec{r}_k$.

Let us write the expansion in a form similar to (46·2), *viz.*,

$$\vec{D(t)} = \sum_{\tau_1, \cdots \tau_r = -\infty}^{+\infty} \vec{D}(\tau_1, \cdots \tau_r) \exp [2\pi i(\tau_1\omega_1 + \cdots + \tau_r\omega_r)t]. \quad (46\cdot6)$$

The sum of the terms of frequency $|\tau_1\omega_1 + \cdots + \tau_r\omega_r|$ is of the form

$$2|\vec{D}(\tau_1, \cdots, \tau_r)| \cos [2\pi t(\tau_1\omega_1 + \cdots + \tau_r\omega_r) + \theta],$$

where θ is a phase constant previously buried in $\vec{D}(\tau_1, \cdots, \tau_r)$. Using this expression for $\vec{d(t)}$ and computing the time average of the power radiated per atom, we obtain

$$I(\tau_1, \cdots, \tau_r) = \frac{64\pi^4}{3c^3}|\tau_1\omega_1 + \cdots + \tau_r\omega_r|^4 \; |\vec{D}(\tau_1, \cdots, \tau_r)|^2. \quad (46\cdot7)$$

Approximate estimates of the intensities of spectrum lines were obtainable on the basis of the Bohr theory by means of (46·7), replacing

$$|\tau_1\omega_1 + \cdots + \tau_r\omega_r|$$

by the actual quantum frequency ν of the emitted radiation and

$$|\vec{D}(\tau_1, \cdots, \tau_r)|^2$$

by some sort of average of its values for the initial and final states under consideration.

The Heisenberg-Born-Jordan matrix theory was a synthesis based on the correspondence principle, the Rydberg combination principle as expressed by (46·1), and the philosophical "hunch" that the material to be used in constructing a proper theory should be more nearly experimentally observable than the hypothetical orbits of the Bohr theory. In the position and intensity of each spectrum line the correspondence principle saw a reflection of the frequency and amplitude of a corresponding harmonic component of the motion. Heisenberg therefore sought to construct a theory which should relate these harmonic components directly to the Hamiltonian function of the atom. The identification of these components requires a double system of indices like that used for matrix elements and the attempt to form algebraic combinations in order to relate the motion to the energy without violating the Rydberg rule led to matrix multiplication and addition. We need not follow the reasoning in detail. Suffice it to say here that from the nature of its,

premises the H. B. J. theory inevitably replaced (46·7) by the formula[1]

$$I(n',n'') = \frac{64\pi^4}{3c^3} \nu_{n'n''}^4 |\vec{D}(n',n'')|^2, \tag{46·8}$$

in which n' and n'' denote the complete sets of quantum numbers characteristic of the initial and final states, respectively, and $\vec{D}(n',n'')$ is the corresponding element of the Heisenberg matrix $\vec{D}$. The formula (46·8) gives the rate of emission of energy per atom when the external radiation field is negligible. The rates of absorption and forced emission in a natural radiation field can be worked out either by the aid of Einstein's quasi-thermodynamic[2] theory of the relation between emission and absorption transition probabilities (*cf.* Sec. 54), or from Van Vleck's formulation of the correspondence principle for absorption.[3]

Thus the Heisenberg theory was, so to speak, born with formulas for the rates of emission and absorption tied around its neck. This scheme for adapting classical intensity formulas to quantum mechanics is entirely successful in outcome but smacks too much of analogy to be entirely satisfying. In Sec. 54 we shall consider the justification of (46·8) from a more advanced point of view.

A serious difficulty with the original Bohr theory lay in the fact that it was applicable only to multiply-periodic motions, whereas a strict application of classical mechanics to the accepted Rutherford atomic model would lead in practically all cases to essentially aperiodic motions. Nevertheless Bohr found it possible to correlate the various empirical quantum jumps accompanying the emission of radiation with corresponding terms of a Fourier's series, using a plausible multiply-periodic *idealization* of the approved model. The formula (46·8) can be established on the basis of quantum mechanics independently of the correspondence principle and is not subject to the above limitation to a special type of motion. On the other hand, it is still sometimes convenient to use a multiply-periodic model as a basis for the approximate estimate of intensities by the correspondence-principle method.

It follows from the relation established in Sec. 11 between the Schrödinger equation and the corresponding classical Hamilton-Jacobi equation that if the variables can be separated in the former by the introduction of proper coordinates they can be separated in the latter by using the same coordinates. If the variables are separable in the Hamilton-

---

[1] In the usual case where the initial and final energy levels are degenerate, one uses (46·8) to compute the intensity of emission from each substate of the upper energy level to each substate of the lower level, adding these contributions to get the total intensity of the line.

[2] A. EINSTEIN, *Phys. Zeits.* **18**, 121 (1917).

[3] J. H. VAN VLECK, *Phys. Rev.* **24**, 330, 347 (1924).

Jacobi equation and if each coordinate has a definite frequency (*cf.* M. Born, *Vorlesungen über Atommechanik*, §14, Berlin, 1925), the classical motion is multiply-periodic. Other types of multiply-periodic motions occur rarely, if at all. Thus it turns out that in practice the correspondence principle is usually applicable only if the variables in the Schrödinger equation are separable.

CHAPTER XI

THEORY OF PERTURBATIONS WHICH DO NOT INVOLVE THE TIME

47. THE PERTURBATION THEORY FOR NONDEGENERATE PROBLEMS

47a. First-order Perturbations.— In quantum mechanics, as in the Bohr theory, perturbation methods are of fundamental importance due to the fact that so few problems can be rigorously solved by direct attack. The essential feature of these methods is that one starts with an approximate solution of the problem in hand and proceeds to compute by "hammer and tongs" a series of corrections designed to improve the approximation. The success of such a computation depends partly on the patience and energy of the computor and partly on his ability to find a happy starting point. The successive approximations may not converge on an exact solution of the problem, but usually the early corrections do yield an appreciable improvement on the initial wave functions.

The conventional perturbation theory of wave mechanics[1] is concerned with the determination of the discrete eigenvalues and eigenfunctions of a Hamiltonian operator $H(q, \partial/\partial q)$. It attempts to approximate these values and functions with the aid of the rigorous solutions of the eigenvalue-eigenfunction problem of a simplified, but related, Hamiltonian function $H_0(q, \partial/\partial q)$ involving the same coordinates. We designate the problems based on the two operators H_0 and H as the unperturbed and the perturbed problems, respectively. In order to pass from the known discrete solutions, say $\psi_k^{(0)}(q)$, $E_k^{(0)}$, of the unperturbed problem to corresponding solutions of the perturbed problem, it is useful to construct a one-parameter continuous series of problems that bridge the gap between the two. To do this we introduce the symbol H_1 for the difference $H - H_0$ and define the operator $\tilde{H}$ by the equation

$$\tilde{H} = H_0 + \lambda H_1 = H_0 + \lambda(H - H_0). \qquad \cdot \quad (47 \cdot 1)$$

Here λ is a parameter which may take on any value between zero and unity. Since $\tilde{H}$ reduces to H_0 when λ is zero and to H when λ is unity, we call $\tilde{H}$ the *interpolation Hamiltonian*. H_1 is called the *perturbing Hamiltonian*. The method of successive approximations which we here apply to $\tilde{H}$ can also be applied to operators expressible as complete power series in a parameter λ.

[1] *Cf.* E. SCHRÖDINGER, *Ann. d. Physik* (4) **80**, 437 (1926).

380

We assume that the solutions of the interpolation problem

$$\tilde{H}\tilde{\psi} = \tilde{E}\tilde{\psi} \tag{47.2}$$

are analytic functions of λ in the interval $0 \leqslant \lambda \leqslant 1$ and therefore reduce to solutions of the unperturbed problem when λ is zero and become solutions of the perturbed problem when λ is unity.

Exceptions to this rule can occur when the perturbing Hamiltonian H_1 introduces new singularities into the problem or modifies in a fundamental way the conditions at preexisting singularities. It can happen, for example, that H_0 has a discrete spectrum while H and $\tilde{H}$ have purely continuous spectra. In such cases solutions of the interpolation problem will change discontinuously when we pass from any positive value of λ to the value zero. *The possibility of such discontinuities requires special examination for each individual case, but in developing the general theory we shall assume that solutions of Eq. (47.2) form a truly continuous connection between the eigenfunctions—and eigenvalues—of the perturbed and unperturbed problems.*

In some cases the interpolated series of problems is wholly artificial, but in others, *e.g.*, when H_1 represents the perturbing effect of an external electric or magnetic field, all values of λ give results of experimental significance and it is unnecessary to distinguish between the interpolation Hamiltonian and the perturbed Hamiltonian. Of course the eventual artificiality of the $\tilde{H}$ problem has nothing to do with the mathematical procedure, or with its validity.

Since degeneracy introduces considerable complications into perturbation theory, we begin with a study of cases in which H_0 and $\tilde{H}$ have nondegenerate eigenvalues.[1] Let there be given a complete normalized set of discrete and continuous-spectrum eigenfunctions $\psi_k^{(0)}(q)$, $\psi_E^{(0)}(q)$ of the unperturbed problem

$$H_0\psi^{(0)} = E^{(0)}\psi^{(0)}. \tag{47.3}$$

Let $\tilde{\psi}_k(\lambda, q)$ denote the kth normalized discrete eigenfunction of the interpolation problem (47.2). If the solutions of the two problems are suitably numbered, our continuity hypothesis requires that

$$\lim_{\lambda \to 0} \tilde{E}_k(\lambda) = E_k^{(0)},$$

and that $\tilde{\psi}_k(0, q) = \lim_{\lambda \to 0} \tilde{\psi}_k(\lambda, q)$ shall be an eigenfunction of (47.3). Then

---

[1] Cases of near degeneracy where the unperturbed eigenvalues occur in closely spaced groups are best treated by a suitable modification of the variation method (Sec. 51), or they can be reduced to degenerate form by pulling out of H_0 the term responsible for the splitting of the groups and adding this term to the perturbing Hamiltonian.

$\tilde{\psi}_k(0,q) = c^{i\alpha}\psi_k^{(0)}(q)$, where α is an arbitrary phase factor which we can set equal to zero without loss of generality.

Expanding $\tilde{E}_k$ and $\tilde{\psi}_k$ into power series in λ, we have

$$\left. \begin{aligned} \tilde{E}_k &= E_k^{(0)} + \sum_{\tau=1}^{\infty} \frac{\lambda^{\tau}}{\tau!}\left(\frac{\partial^{\tau}\tilde{E}_k}{\partial\lambda^{\tau}}\right)_{\lambda=0} \\ \tilde{\psi}_k &= \psi_k^{(0)} + \sum_{\tau=1}^{\infty} \frac{\lambda^{\tau}}{\tau!}\left(\frac{\partial^{\tau}\tilde{\psi}_k}{\partial\lambda^{\tau}}\right)_{\lambda=0} \end{aligned} \right\} \tag{47.4}$$

Introducing the abbreviations

$$\left. \begin{aligned} E_k^{(\tau)} &= \frac{1}{\tau!}\left(\frac{\partial^{\tau}\tilde{E}_k}{\partial\lambda^{\tau}}\right)_{\lambda=0}, \\ \psi_k^{(\tau)} &= \frac{1}{\tau!}\left(\frac{\partial^{\tau}\tilde{\psi}_k}{\partial\lambda^{\tau}}\right)_{\lambda=0} \end{aligned} \right\} \tag{47.5}$$

we reduce these series to the form

$$\left. \begin{aligned} \tilde{E}_k &= E_k^{(0)} + \lambda E_k^{(1)} + \lambda^2 E_k^{(2)} + \cdots, \\ \tilde{\psi}_k &= \psi_k^{(0)} + \lambda\psi_k^{(1)} + \lambda^2\psi_k^{(2)} + \cdots \end{aligned} \right\} \tag{47.6}$$

Our task is now to compute enough terms of these two series to obtain satisfactory approximations to the perturbed E_k and ψ_k. Whether or not the series actually converge for $\lambda = 1$ to give a rigorous solution of the perturbed problem is a question we shall not attempt to discuss. If the singular-point characteristics of H_0 and H are the same, there can be no reasonable doubt that these series will converge[1] for sufficiently small values of λ, and if the initial approximation is good, we may hope not only for convergence but for rapid convergence when λ is unity. Unfortunately it is usually impracticable to compute more than two or three terms of these series so that the usefulness of the method is limited to cases in which the zero-order approximations $E_k^{(0)}$, $\psi_k^{(0)}$ are quite good.

Let us now develop a scheme for computing the *first-order corrections* $E_k^{(1)}$, $\psi_k^{(1)}$ to the energy and wave function, respectively. Using the notation of Eq. (47.5), differentiating (47.2) with respect to λ, and giving λ the value zero in the resulting equation, we obtain

$$(H_0 - E_k^{(0)})\psi_k^{(1)} = (E_k^{(1)} - H_1)\psi_k^{(0)}. \tag{47.7}$$

This equation can be used for the determination of both $E_k^{(1)}$ and $\psi_k^{(1)}$. It is not difficult to show that the left-hand member of this equation is

[1] A. H. Wilson makes the statement that, at least in cases which involve only discrete spectra, the convergence is that of an exponential series if there is any convergence at all. [*Cf.* A. H. Wilson, *Proc. Roy. Soc.* A, **124**, 186 (1929).]

orthogonal to $\psi_k^{(0)}$. For, since H_0 is Hermitian with respect to functions of class D, it follows from (47·3) that

$$\int_\infty \psi_k^{*(0)}(H_0 - E_k^{(0)})\psi_k^{(1)}d\tau = \int_\infty \psi_k^{(1)}(H_0^* - E_k^{(0)})\psi_k^{*(0)}d\tau = 0. \quad (47\cdot8)$$

Hence the right-hand member of (47·7) is also orthogonal to $\psi_k^{(0)}$ and the value of $E_k^{(1)}$ is fixed by

$$E_k^{(1)} = \int_\infty \psi_k^{*(0)}H_1\psi_k^{(0)}d\tau = (H_1\psi_k^{(0)}, \psi_k^{(0)}). \quad (47\cdot9)$$

In words Eq. (47·9) states that the *first-order energy correction is equal to the mean value of the perturbing energy operator averaged over the unperturbed, or zero-order, wave function* [cf. Eq. (35·4)]. This is the quantum-mechanical equivalent of a familiar theorem of the perturbation theory of classical mechanics.[1]

We may now regard Eq. (47·7) as an inhomogeneous equation for the partial determination of $\psi_k^{(1)}$. The unknown function is not fully determined by this equation, for we can add to any solution an arbitrary multiple of $\psi_k^{(0)}$ and get a new solution as good as the first. This indeterminateness in $\psi_k^{(1)}$ is partially removed when we take into account the normalization requirement for ψ_k, which yields

$$\left[\frac{\partial}{\partial\lambda}\int_\infty \psi_k\psi_k^*d\tau\right]_{\lambda=0} \equiv \int_\infty \psi_k^{(0)}\psi_k^{*(1)}d\tau + \int_\infty \psi_k^{*(0)}\psi_k^{(1)}d\tau = 0.$$

Hence the real part of the scalar product of $\psi_k^{(0)}$ and $\psi_k^{(1)}$ is zero. To fix $\psi_k^{(1)}$ uniquely we arbitrarily agree that the imaginary part of this scalar product shall vanish also. Thus

$$(\psi_k^{(0)}, \psi_k^{(1)}) \equiv \int_\infty \psi_k^{(0)}\left(\frac{\partial\psi_k^*}{\partial\lambda}\right)_{\lambda=0}d\tau = 0. \quad (47\cdot10)$$

In order to solve (47·7) we expand the unknown function $\psi_k^{(1)}$ into a series of eigenfunctions of the unperturbed problem. In view of Eq. (47·10) there will be no term in $\psi_k^{(0)}$. The expansion takes the form

$$\psi_k^{(1)} = \sum_{n\neq k}\psi_n^{(0)}U^{(1)}(n,k) + \int\psi_E^{(0)}U^{(1)}(E,k)dE, \quad (47\cdot11)$$

where the coefficients $U^{(1)}(n,k)$, $U^{(1)}(E,k)$ are unknowns to be determined. The scheme is now to express each side of Eq. (47·7) as a linear combination of the known unperturbed eigenfunctions and to equate coefficients of corresponding terms. For this purpose we need the expansion of $H_1\psi_k^{(0)}$, which we write in the form

$$H_1\psi_k^{(0)} = \sum_n\psi_n^{(0)}H_1(n,k) + \int\psi_E^{(0)}H_1(E,k)dE. \quad (47\cdot12)$$

---

[1] *Cf.*, for instance, J. H. VAN VLECK, *Quantum Principles and Line Spectra*, p. 203, Washington, 1926; or M. BORN, *Vorlesungen über Atommechanik*, p. 287, Berlin, 1925.

Assuming the usual normalization for the discrete functions $\psi_n^{(0)}$ the matrix components $H_1(n,k)$ are fixed in accordance with (44·2) by the explicit formula

$$H_1(n,k) = \int \psi_n^{*(0)} H_1 \psi_k^{(0)} d\tau. \tag{47·13}$$

Similarly if $H_1 \psi_k^{(0)}$ is absolutely integrable [*cf.* Eqs. (36·13) and (36·14) and remark following latter equation],

$$H_1(E,k) = \int \psi^{*(0)} H_1 \psi_k^{(0)} d\tau. \tag{47·14}$$

Inserting the expansions (47·11) and (47·12) into Eq. (47·7) and equating coefficients of like terms on the two sides,[1] we obtain

$$U^{(1)}(n,k) = \frac{H_1(n,k)}{E_k^{(0)} - E_n^{(0)}}, \qquad n \neq k \tag{47·15}$$

$$U^{(1)}(E,k) = \frac{H_1(E,k)}{E_k^{(0)} - E}. \tag{47·16}$$

These formulas combined with Eq. (47·11) complete the determination of the first-order correction to the wave functions.

47b. Second-order Perturbations.—In order to get the second-order corrections to E_k and ψ_k we differentiate Eq. (47·2) *twice* with respect to λ and then set λ equal to zero. Using the notation of Eqs. (47·5) the result takes the form

$$(H_0 - E_k^{(0)})\psi_k^{(2)} = E_k^{(2)}\psi_k^{(0)} + (E_k^{(1)} - H_1)\psi_k^{(1)}. \tag{47·17}$$

As in the case of Eq. (47·7) the left-hand member is orthogonal to $\psi_k^{(0)}$ and we have

$$E_k^{(2)} = \int_\infty \psi_k^{*(0)}(H_1 - E_k^{(1)})\psi_k^{(1)} d\tau = \int_\infty \psi_k^{*(0)} H_1 \psi_k^{(1)} d\tau. \tag{47·18}$$

In virtue of Eqs. (47·11), (47·15), and (47·16) the expression for the second-order energy correction takes the final form

$$E_k^{(2)} = \sum_{n \neq k} \frac{|H_1(k,n)|^2}{E_k^{(0)} - E_n^{(0)}} + \int \frac{|H_1(k,E)|^2}{E_k^{(0)} - E} dE. \tag{47·19}$$

It will be observed that in this approximation each energy level is pushed upward by those below it and downward by those above it. The displacements are proportional to the squares of the corresponding matrix elements of H_1 and inversely proportional to the separations of the unperturbed energy levels involved. Thus in general each level is most affected by its near neighbors. This tendency is accentuated by

[1] The equating of coefficients can be justified by taking the scalar product of each side of the equation and the successive discrete eigenfunctions and eigendifferentials of H_0.

the fact that the matrix elements $H_1(k,n)$, $H_1(k,E)$ are usually small when the corresponding energy-level differences are large.

In the case of the lowest energy level the quantity $E_0^{(2)}$ is essentially negative so that the graph of $\bar{E}_0$ against λ is curved downward at the point $\lambda = 0$. The same argument can be applied for any other value of λ and we infer that the graph in question is concave downward throughout its course.

The attention of the reader is called to the fact that to compute $E_k^{(1)}$ and $E_k^{(2)}$ we need only evaluate the approximations of *next lower order* to the wave function. This is a general characteristic of perturbation calculations in the semiclassical Bohr theory[1] as well as in quantum mechanics. It means that for a given amount of labor we can always compute the energy levels more accurately than the wave functions. In this connection it is illuminating to recall that eigenvalues of the energy are the extreme values of integrals over the wave functions. Thus any small error in the assumed wave function tends to produce an error of higher order—*i.e.*, a smaller error—in the energy.

Consider next the second-order correction to the wave function, which we represent by the series-integral expansion

$$\psi_k^{(2)} = \sum_n{}' \psi_n^{(0)} U^{(2)}(n,k) + \int \psi_E^{(0)} U^{(2)}(E,k)dE \qquad (47 \cdot 20)$$

with unknown coefficients. A similar expansion is assumed for the function $H_1 \psi_k^{(1)}$:

$$H_1 \psi_k^{(1)} = \sum_n{}' \psi_n^{(0)} h_{nk} + \int \psi_E^{(0)} h_{Ek} dE, \qquad (47 \cdot 21)$$

where

$$h_{nk} = \int_\infty \psi_n^{*(0)} H_1 \Big[\sum_m \psi_m^{(0)} U^{(1)}(m,k) + \int \psi_E^{(0)} U^{(1)}(E,k)dE \Big] d\tau$$

$$= \sum_{m \neq k} \frac{H_1(n,m)H_1(m,k)}{E_k^{(0)} - E_m^{(0)}} + \int \frac{H_1(n,E)H_1(E,k)}{E_k^{(0)} - E} dE, \qquad (47 \cdot 22)$$

$$h_{Ek} = \int_\infty \psi_E^{*(0)} H_1 \Big[\sum_m \psi_m^{(0)} U^{(1)}(m,k) + \int \psi_E^{(0)} U^{(1)}(E,k)dE \Big] d\tau. \qquad (47 \cdot 23)$$

Each side of Eq. (47·17) can now be written as a linear combination of unperturbed eigenfunctions. Equating corresponding coefficients on the two sides of the resulting equation, we obtain

$$0 = E_k^{(2)} - h_{kk}, \qquad (47 \cdot 24)$$

$$(E_n^{(0)} - E)_k^{(0)} U^{(2)}(n,k) = E_k^{(1)} U^{(1)}(n,k) - h_{nk}, \qquad n \neq k \quad (47 \cdot 25)$$

$$(E - E_k^{(0)}) U^{(2)}(E,k) = E_k^{(1)} U^{(1)}(E,k) - h_{Ek}. \qquad (47 \cdot 26)$$

[1] *Cf.* references in footnote 1, p. 383.

The first of these equations is equivalent to Eq. (47·19), while Eq. (47·25) in conjunction with (47·22) yields

$$U^{(2)}(n,k) = \frac{1}{E_k^{(0)} - E_n^{(0)}} \left[\sum_{m \neq k} \frac{H_1(n,m)H_1(m,k)}{E_k^{(0)} - E_m^{(0)}} + \int \frac{H_1(n,E)H_1(E,k)}{E_k^{(0)} - E} dE \right.$$
$$\left. - \frac{H_1(n,k)H_1(k,k)}{E_k^{(0)} - E_n^{(0)}} \right]. \qquad n \neq k \quad (47·27)$$

To avoid undue complications we leave the expression for $U^{(2)}(E,k)$ in the form

$$U^{(2)}(E,k) = \frac{1}{E_k^{(0)} - E} \left[h_{Ek} - \frac{H_1(E,k)H_1(k,k)}{E_k^{(0)} - E} \right]. \qquad (47·28)$$

The coefficient $U^{(2)}(k,k)$ is not determined by Eq. (47·27) but is restricted by the normalization condition for ψ_k. Differentiating the equation which expresses this condition twice with respect to λ and setting λ equal to zero, we obtain, in analogy with Eq. (47·10), the relation

$$\int_\infty (\psi_k^{*(0)}\psi_k^{(2)} + \psi_k^{(0)}\psi_k^{*(2)})d\tau = -\int_\infty |\psi_k^{(1)}|^2 d\tau. \qquad (47·29)$$

Hence

$$U^{(2)}(k,k) + U^{(2)}(k,k)^* = -\left[\sum_n |U^{(1)}(n,k)|^2 + \int |U^{(1)}(E,k)|^2 dE \right].$$

We adopt the simplest choice of $U^{(2)}(k,k)$ consistent with the above relation, *viz.*,

$$U^{(2)}(k,k) = -\tfrac{1}{2}\left[\sum_n |U^{(1)}(n,k)|^2 + \int |U^{(1)}(E,k)|^2 dE \right]. \quad (47·30)$$

Formulas have now been developed for the determination of all the coefficients in Eq. (47·20) and our study of the second-order correction to the wave function is ·complete. These second-order formulas are so complicated that they are seldom used and the corrections of the third and higher order are still more complex.[1] Hence we carry the work no farther.

47c. An Example: The Diatomic Molecule.—We consider here only a very simple and elementary application of the above perturbation formulas to the dumbbell model of the diatomic molecule discussed in Sec. 28.

The radial differential equation (28·19) can be written in the form

$$H\mathfrak{R} = \left[-\frac{h^2}{8\pi^2\mu} \frac{d^2}{dr^2} + V(r) + \frac{l(l+1)h^2}{8\pi^2\mu r^2} \right]\mathfrak{R} = E\mathfrak{R}. \qquad (47·31)$$

[1] *Cf.*, however, K. F. NIESSEN, *Phys. Rev.* **34**, 263 (1929) for energy formulas of third and fourth order.

The "centrifugal-force" term in this equation can be treated as a perturbation on the simplified equation

$$H_0 \Re^{(0)} = \left[-\frac{h^2}{8\pi^2\mu} \frac{d^2}{dr^2} + V(r) \right] \Re^{(0)} = E^{(0)} \Re^{(0)}. \qquad (47\cdot32)$$

If the minimum of $V(r)$ is deep, the lower energy levels of the unperturbed problem will be spaced at approximately equal intervals like those of an ideal linear oscillator and the corresponding wave functions will resemble those of a linear oscillator of suitable frequency with potential minimum at $r = r_0$.

Before applying perturbation methods, however, we must see whether they are really legitimate in this case. To this end we observe in the first place that the perturbing Hamiltonian operator

$$H_1 = l(l+1)\frac{h^2}{8\pi^2\mu r^2} \times$$

introduces no new singular points into the problem. It vanishes rapidly at infinity, thus insuring that the perturbed wave functions will behave like the unperturbed at the outer singular point. The fact that H_1 becomes infinite more rapidly than $V(r)$ at $r = 0$ looks suspicious but leads to no real difficulty. It follows from the work of Sec. 28 that solutions of the unperturbed problem vanish as r to the first power at the origin while those of the perturbed problem vanish as r^{l+1}. It is easily shown that eigenfunctions of the interpolation problem vanish at the origin as a power γ of r which reduces to unity, or $l+1$, when λ is given the values 0 and 1, respectively. Thus the solutions of the interpolation problem pass continuously into solutions of the unperturbed or perturbed problems at the ends of the range $0 \leqslant \lambda \leqslant 1$. Finally the integrals which form the matrix elements of H_1 exist despite the pole of H_1 at the origin.

We can now apply the elementary formula (47·9), denoting the mean value of the reciprocal of the moment of inertia μr^2 over the vth unperturbed radial eigenfunction by $1/I_v$. To first-order corrections the energy of the model is

$$E = E(v,l) = E_v^{(0)} + \frac{l(l+1)h^2}{8\pi^2 I_v}. \qquad \left\{ \begin{array}{l} l = 0, 1, 2, \cdots \\ v = 0, 1, 2, \cdots \end{array} \right\} \qquad (47\cdot33)$$

This equation is seen to be in harmony with the approximate result (28·24) previously obtained by the method of Brillouin, Wentzel, and Kramers if we identify I_v with μr_0^2 and note that the difference between the rotational terms $\left(l + \frac{1}{2}\right)^2 \frac{h^2}{8\pi^2\mu r_0^2}$ and $\frac{l(l+1)h^2}{8\pi^2 I_v}$ is a constant which can be absorbed into $E_v^{(0)}$. The perturbation-theory method here

used is the simplest method of deriving the rotational term, but on the whole the B. W. K. method when used in higher approximation is perhaps the most satisfactory for dealing with this molecular problem.[1]

48. THE PERTURBATION THEORY FOR DEGENERATE PROBLEMS

48a. First-order Energy Perturbations.—The straightforward perturbation theory for initially degenerate systems has a tendency to become complicated. The various schemes for reducing the complexities are best appreciated, however, after an introduction to the problem
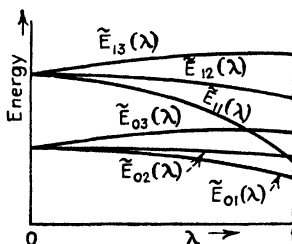
—such as that which follows—which makes no use of tricks. Discussion of the perturbation theory for degenerate problems is facilitated by reference to a graph of the interpolation energies $\bar{E}(\lambda)$ plotted against λ such as that shown in Fig. 19. The degree of degeneracy or statistical weight g of an energy level being defined as the number of linearly independent eigenfunctions for the given energy value, it is evident *a priori* that this quantity must be independent of λ except at points where two or more energy curves meet or cross. At such points the statistical weight is equal to the sum of its values for the different curves which meet. If the symmetry of the perturbed problem is the same as that of the unperturbed, the degeneracy of each level of the interpolation problem will usually be the same as that of a corresponding unperturbed level. In this case the curves for the different interpolation levels will have no tendency to meet at the $\lambda = 0$ axis. Very frequently, however, the perturbed and interpolation problems are less symmetrical than the unperturbed, and any given unperturbed energy value $E_k^{(0)}$ can be the starting point of several divergent $\bar{E}, \lambda$ curves. In such cases an *arbitrary* unperturbed wave function $u(q)$ is not necessarily the limit as λ approaches zero of some interpolation function $\bar{\psi}(\lambda, q)$. Thus if $\bar{\psi}_a$ and $\bar{\psi}_b$ are two interpolation functions whose energy curves meet at $\lambda = 0$, it is evident that any linear combination of $\bar{\psi}_a(0,q)$ and $\bar{\psi}_b(0,q)$ makes a possible unperturbed function although it will not be the limit of any $\bar{\psi}(\lambda,q)$. Hence the first term in the expansion of $\bar{\psi}$ in powers of λ cannot be assumed to be known even if the unperturbed problem has been fully treated. The determination of this first term, or *zero-order approximation*, is a process which must be carried along with the determination of the energy and the higher order terms in $\bar{\psi}$.

[1] *Cf.* E. Fues, *Ann. d. Phys.* **80**, 367 (1926) for the classical application of the perturbation-theory method to this problem. The most complete treatment by means of the B. W. K. method has been given by J. L. Dunham (*cf.* footnote 1, p. 157).

Let us now suppose that we have given a complete orthonormal system of eigenfunctions of the unperturbed-problem Hamiltonian H_0. Each discrete function will be denoted by a symbol u_{ki}, where the first subscript, k, indicates the energy level, while i indicates the individual member of the group of eigenfunctions which belong to this level. According to Sec. 32k the discrete unperturbed levels have a finite degeneracy, so that the index i ranges from 1 to an upper limit g_k which is the statistical weight of the energy level $E_k^{(0)}$. (To avoid an undue multiplicity of subscripts we shall at times omit the subscript on the symbol g when no ambiguity will result.) The continuous-spectrum wave functions must be provided with similar double subscripts E and i. Finally the interpolation wave functions will also be specified by two subscripts k, i of which the first indicates the parent unperturbed energy level.

Equation (47·7) is applicable in the present case provided that additional subscripts are introduced. Thus

$$(H_0 - E_k^{(0)})\psi_{ki}^{(1)} = (E_{ki}^{(1)} - H_1)\psi_{ki}^{(0)}. \tag{48·1}$$

By the argument of the preceding section the left-hand member is orthogonal to all eigenfunctions of the unperturbed equation belonging to the energy level $E_k^{(0)}$. Hence it is orthogonal to $u_{k1}, \cdots, u_{kg}$ and $\psi_{k1}^{(0)}, \cdots, \psi_{kg}^{(0)}$. The right-hand member must have the same properties. Forming the scalar product of the right-hand member with $\psi_{ki}^{(0)}$, we obtain

$$E_{ki}^{(1)} = (H_1\psi_{ki}^{(0)}, \psi_{ki}^{(0)}) \tag{48·2}$$

in analogy with (47·9). In this case, however, the equation is not so directly useful since the functions $\psi_{ki}^{(0)}$ are as yet unknown.

The functions $\psi_{ki}^{(0)}$ are of course expressible in the form

$$\psi_{ki}^{(0)} = \sum_j u_{kj} U^{(0)}(kj; ki). \tag{48·3}$$

If the interpolation functions $\bar{\psi}_{ki}$ form an orthonormal set, as we shall suppose, the zero-order functions $\psi_{ki}^{(0)}$ must do the same. Under these circumstances the coefficients $U^{(0)}(kj; ki)$ of (48·3) are subject to the equations

$$\left. \begin{aligned} \sum_{j=1}^{g} U^{(0)}(kj; ki) U^{(0)}(kj; ki')^* &= \delta_{ii'}; \\ \sum_{i=1}^{g} U^{(0)}(kj; ki) U^{(0)}(kj'; ki)^* &= \delta_{jj'}. \end{aligned} \right\} \tag{48·4}$$

In fact the set of $U^{(0)}$ coefficients for each unperturbed energy level form a finite unitary matrix and by adding zero elements connecting different

unperturbed energies we can form an infinite unitary step matrix $U^{(0)}$ like the matrix V of p. 363. Substituting the sum (48·3) into the right-hand member of Eq. (48·1) and setting the scalar products of the resulting expression and the successive functions u_{kl} $(l = 1,2, \cdots ,g_k)$ in turn equal to zero we obtain the set of g_k equations

$$\sum_{j=1}^{g_k}[H_1(kj';kj) - E_{ki}^{(1)}\delta_{j'j}]U^{(0)}(kj;ki) = 0, \qquad j' = 1,2,3, \cdots ,g_k \quad (48\cdot5)$$

in which $H_1(kj';kj)$ is the matrix element

$$H_1(kj';kj) = \int_\infty u_{kj'}{}^*H_1u_{kj}d\tau = (H_1u_{kj},u_{kj'}). \quad (48\cdot6)$$

The above set of g_k simultaneous linear homogeneous equations in the g_k unknowns $U^{(0)}(k1;ki)$, $U^{(0)}(k2;ki)$, $\cdots$, $U^{(0)}(kg_k;ki)$ is formally the
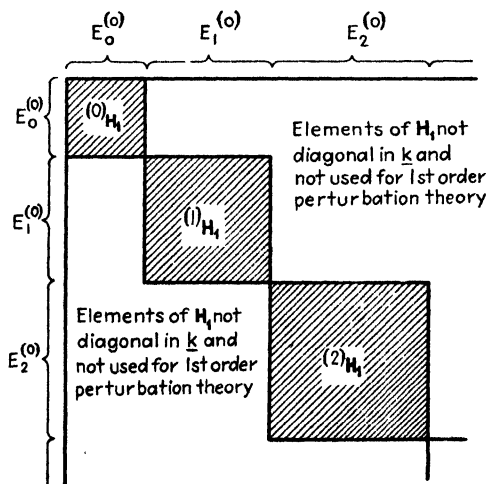
FIG. 20.—Diagram of the steps of H_1 initially diagonal in k and to be completely diagonal-ized by canonical transformation with matrix $U^{(0)}$.

same as the set of equations (44·24) or (44·30). Hence the theory of the eigenvalue-eigenfunction problem for finite matrices, as developed in Sec. 44c, is applicable. Nontrivial solutions can exist if, and only if, $E_{ki}^{(1)}$ is a root of the secular equation obtained by setting the deter-minant of the coefficients equal to zero. Let us denote the $g_k \times g_k$ element Hermitian matrix $\|H_1(kj';kj)\|$, in which k is fixed, by $^{(k)}H_1$ ($cf.$ Fig. 20). Then we can write the secular equation in the form

$$\det (^{(k)}H_1 - E_{ki}^{(1)} I) = 0. \quad (48\cdot7)$$

This equation has g_k real roots. If they' are all distinct, there are g_k different initial slopes for the $\bar{E}$ versus λ curves (Fig. 19) which originate at the kth unperturbed level. Hence there must be g_k nondegenerate interpolation levels which unite to form a single degenerate level $E_k^{(0)}$

when λ is zero. At the other extreme is the case where the off-diagonal elements of the $^{(k)}\mathbf{H}_1$ matrix are zero while the diagonal elements are equal. In this case, (48·7) has a single g_k-fold root and the interpolation problem in this approximation has the same degeneracy as the unperturbed problem.

The set of g_k unknowns $U^{(0)}(kj;ki)$ for each value of i constitutes a vector $\vec{\xi}_i$ like the vectors $\vec{\xi}$ and $\vec{x}$ of Sec. 44c. It can always be normalized and will be automatically orthogonal to any similar vector derived from a different root of the secular equation. Whether the roots are distinct or multiple it is always possible to choose a set of eigenvectors, including p for each p-fold root, which are mutually orthogonal and unite to form a $g_k \times g_k$ element unitary matrix $\| U^{(0)}(kj;ki) \|$.

48b. Second-order Energy Perturbations.—Our first-order energy and zero-order wave function calculation has led us to a simple matrix eigenvalue-eigenvector problem. Before proceeding farther it will be useful to restate the entire perturbation problem in matrix form. For this purpose we identify the operator α of (44·23) with the Hamiltonian operator H. Let us at the same time generalize our previous treatment by assuming that $\tilde{H}$ is not necessarily linear in λ but is representable by a power series. For our basic set of φ functions we take the unperturbed wave functions u_{ki}. We write the corresponding initial matrix as $\tilde{\mathbf{H}} = \mathbf{H}_0 + \lambda \mathbf{H}_1 + \lambda^2 \mathbf{H}_2 + \cdots$ and take $\mathbf{H}_0$ to be an ordered diagonal matrix. For the diagonal matrix $\mathbf{A}$ of Eq. (44·33) to be obtained by the transformation of $\tilde{\mathbf{H}}$ we introduce the symbol $\tilde{\mathbf{E}}$ since it is built up from the interpolation energy values. Thus (44·33) becomes

$$(\mathbf{H}_0 + \lambda \mathbf{H}_1 + \lambda^2 \mathbf{H}_2 + \cdots)\tilde{\mathbf{U}} = \tilde{\mathbf{U}}\tilde{\mathbf{E}}. \tag{48·8}$$

Following the procedure of Sec. 47 we seek a solution in power series:

$$\left. \begin{aligned} \tilde{\mathbf{U}} &= \mathbf{U}^{(0)} + \lambda \mathbf{U}^{(1)} + \lambda^2 \mathbf{U}^{(2)} + \cdots, \\ \tilde{\mathbf{E}} &= \mathbf{E}^{(0)} + \lambda \mathbf{E}^{(1)} + \lambda^2 \mathbf{E}^{(2)} + \cdots. \end{aligned} \right\} \tag{48·9}$$

Here $\mathbf{E}^{(0)}, \mathbf{E}^{(1)}, \mathbf{E}^{(2)}, \cdots$ are diagonal and $\mathbf{E}^{(0)}$ must contain the same diagonal elements as $\mathbf{H}_0$. We assume that $\mathbf{E}^{(0)}$ is identical with $\mathbf{H}_0$. This means that the numbering of the $\tilde{\psi}$'s is made to harmonize with that of the u_{ki}'s. Inserting the expansion into (48·8) and equating the coefficients of like powers of λ to zero, we obtain the sequence of equations

$$\mathbf{E}^{(0)}\mathbf{U}^{(0)} - \mathbf{U}^{(0)}\mathbf{E}^{(0)} = 0, \tag{48·10}$$

$$\mathbf{E}^{(0)}\mathbf{U}^{(1)} - \mathbf{U}^{(1)}\mathbf{E}^{(0)} = \mathbf{U}^{(0)}\mathbf{E}^{(1)} - \mathbf{H}_1\mathbf{U}^{(0)}, \tag{48·11}$$

$$\mathbf{E}^{(0)}\mathbf{U}^{(2)} - \mathbf{U}^{(2)}\mathbf{E}^{(0)} = \mathbf{U}^{(0)}\mathbf{E}^{(2)} + \mathbf{U}^{(1)}\mathbf{E}^{(1)} - \mathbf{H}_1\mathbf{U}^{(1)} - \mathbf{H}_2\mathbf{U}^{(0)}, \tag{48·12}$$

. . . .

If a unitary matrix $\tilde{\mathbf{U}}$ can be found which satisfies these equations, it will transform the u_{ki} functions into the $\tilde{\psi}_{ki}$'s; $\mathbf{U}^{(0)}$ will transform the u_{ki}'s into the $\psi_{ki}^{(0)}$'s, etc.

In view of the unitary character of $\tilde{\mathbf{U}}$ it is necessary that the equations

$$\mathbf{U}^{(0)}\mathbf{U}^{(0)\dagger} = \mathbf{I}, \tag{48.13}$$

$$\mathbf{U}^{(1)}\mathbf{U}^{(0)\dagger} + \mathbf{U}^{(0)}\mathbf{U}^{(1)\dagger} = 0, \tag{48.14}$$

$$\mathbf{U}^{(2)}\mathbf{U}^{(0)\dagger} + \mathbf{U}^{(0)}\mathbf{U}^{(2)\dagger} + \mathbf{U}^{(1)}\mathbf{U}^{(1)\dagger} = 0, \tag{48.15}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,$$

shall also be satisfied.

Equations (48.10) and (48.13) reaffirm the conclusion reached on p. 391 that $\mathbf{U}^{(0)}$ is a unitary step matrix with no nonvanishing elements connecting different unperturbed energy levels. Taking the $(kj;ki)$ element of each side of (48.11) for values of j ranging from 1 to g_k, we obtain the simultaneous equations (48.5) and thereby rederive the conclusion that $\mathbf{U}^{(0)}$ must diagonalize the step matrix obtained from $\mathbf{H}_1$ by striking out all elements connecting different unperturbed energies.

The determination of $\mathbf{U}^{(1)}$ and $\mathbf{E}^{(2)}$ with the aid of (48.12) and the remaining elements of (48.11) leads to algebraic difficulties unless the minors $^{(k)}\mathbf{H}_1$ happen to be diagonal. Hence it is convenient to apply to $\tilde{\mathbf{H}}$ a canonical transformation with a matrix $\mathbf{V}$ chosen like $\mathbf{U}^{(0)}$ to diagonalize each $^{(k)}\mathbf{H}_1$. We distinguish between $\mathbf{V}$ and $\mathbf{U}^{(0)}$, because the latter is not yet fully determined unless all the roots of the secular equation (48.7) happen to be distinct. $\mathbf{V}$ is simply an arbitrary solution of (48.10) whose steps $^{(k)}\mathbf{V}$ satisfy the additional equations

$$^{(k)}\mathbf{H}_1{}^{(k)}\mathbf{V} = {}^{(k)}\mathbf{V}\mathbf{E}^{(1)}. \tag{48.16}$$

As a result of the transformation $\mathbf{H}_1$ is replaced by a new matrix

$$\hat{\mathbf{H}}_1 = \mathbf{V}^{-1}\mathbf{H}_1\mathbf{V}$$

with every $^{(k)}\hat{\mathbf{H}}_1$ diagonal. In place of Eqs. (48.8) to (48.15) we have a formally identical set with $\tilde{\mathbf{U}}$ replaced by $\hat{\mathbf{U}} = \mathbf{V}^{-1}\tilde{\mathbf{U}}$. We specify the new set of basic wave functions

$$w_{kl} = \sum_{j=1}^{g_k} u_{kj}V(kj;kl)$$

and the rows and columns of $\hat{\mathbf{H}}_1$ by indices k and l instead of the old indices k and j.

Let $\hat{\mathbf{U}}^{(n)}$ denote the coefficient of λ^n in the expansion of $\hat{\mathbf{U}}$ in powers of λ. In view of the discussion on p. 363 and the fact that both $\mathbf{U}^{(0)}$ and $\mathbf{V}$ diagonalize every $^{(k)}\mathbf{H}_1$, $\hat{\mathbf{U}}^{(0)}$ is readily seen to be a step matrix with no nonvanishing elements connecting different values of $\mathbf{E}^{(0+1)} = \mathbf{E}^{(0)} + \mathbf{E}^{(1)}$. If the unperturbed levels $E_k^{(0)}$ are actually split by the first-order energy calculation, the steps of $\hat{\mathbf{U}}^{(0)}$ will be smaller, in general, than those of $\mathbf{U}^{(0)}$.

Let us now apply the transformed set of Eqs. (48.12) to the determination of the second-order energy perturbation $E_k^{(2)}$. Every $(kl;ki)$

element of the left-hand member vanishes automatically. Setting the corresponding elements of the right-hand member equal to zero, and remembering that every $^{(k)}\hat{\mathbf{H}}_1$ is diagonal, we get

$$\hat{U}^{(0)}(kl;ki)E_{ki}^{(2)} - \hat{U}^{(1)}(kl;ki)(E_{kl}^{(1)} - E_{ki}^{(1)})$$

$$- \sum_{k' \neq k}\sum_{l'}\hat{H}_1(kl;k'l')\hat{U}^{(1)}(k'l';ki) - \sum_{l'}\hat{H}_2(kl;kl')\hat{U}^{(0)}(kl';ki) = 0. \quad (48\cdot17)$$

Let us now choose l and i so that $E_{kl}^{(1)}$ and $E_{ki}^{(1)}$ are equal, thus eliminating the second term of the above equation. For $k' \neq k$, Eq. (48·11) yields

$$\hat{U}^{(1)}(k'l';ki) = \frac{\sum_l \hat{H}_1(k'l';kl)\hat{U}^{(0)}(kl;ki)}{E_k^{(0)} - E_{k'}^{(0)}}.$$

Substitution of this value of $\hat{U}^{(1)}(k'l';ki)$ into (48·17) gives

$$\hat{U}^{(0)}(kl;ki)E_{ki}^{(2)} - \sum_{l''}F(kl;kl'')\hat{U}^{(0)}(kl'';ki) = 0, \qquad (48\cdot18)$$

with $F(kl;kl'')$ defined by

$$F(kl;kl'') = \frac{\sum_{k' \neq k}\sum_{l'}\hat{H}_1(kl;k'l')\hat{H}_1(k'l';kl'')}{E_k^{(0)} - E_{k'}^{(0)}} + \hat{H}_2(kl;kl''). \quad (48\cdot19)$$

On account of the step form of $\hat{\mathbf{U}}^{(0)}$ the l'' sum in (48·18) is to be extended only over one of the small squares corresponding to a single pair of values of $E^{(0)}$ and $E^{(1)}$. For any given i the index l has the same range as l''. Hence the number of Eqs. (48·18) for any given i is equal to the number of unknown elements of $\hat{\mathbf{U}}^{(0)}$ which appear and we can use these equations like (48·5) to determine the energy correction $E_{ki}^{(2)}$. In fact $E_{ki}^{(2)}$ must be a root of the secular equation

$$\det\left(^{(ki)}\mathbf{F} - E^{(2)}\mathbf{I}\right) = 0, \qquad (48\cdot20)$$

where $^{(ki)}\mathbf{F}$ is a diagonal square of $\mathbf{F}$ corresponding to a single value of $E_{ki}^{(1)}$.

Usually the ultimate degeneracy of the perturbed problem is that which remains after the application of the first-order perturbation.[1] When that is the case the eigenvalues of each of the small $^{(ki)}\mathbf{F}$ matrices are all equal. In other words each of these small matrices is transformed by $\hat{\mathbf{U}}^{(0)}$ into a multiple of a unit matrix. It follows that these $\mathbf{F}$ matrices are *initially* multiples of unity and that every $^{(ki)}\mathbf{F}$ is initially diagonal.

---

[1] This will surely be the case if the initial degeneracy is entirely removed by the first-order perturbation. Otherwise, we may know from a study of the integrals of the problem what the ultimate degeneracy is. Then, if the first-order perturbation splits each unperturbed level into the expected final number of sublevels, we can be sure that no further splitting will come from the higher order perturbations.

Under these circumstances (48·18) must reduce to

$$E_{ki}^{(2)} = F(kl;kl) = \frac{\sum_{k' \neq k} \sum_{l'} |\widehat{H}_1(kl;k'l')|^2}{E_k^{(0)} - E_l^{(0)}}, \tag{48·21}$$

and the off-diagonal elements $F(kl;kl'')$ need not be computed. It will be observed that (48·21) is essentially the same as (47·19) except that it is based on $\widehat{H}_1$ instead of H_1.

*48c. Van Vleck's Method for Second-order Perturbations.—Unfortunately, it is often difficult, or impossible, to apply the transformation V to the whole of H_1, although there may be no difficulty in diagonalizing those steps of H_1 which belong to the lower unperturbed energy levels in which there is the greatest physical interest. This difficulty is the more serious because the upper, continuous portion of the energy spectrum can be brought within the scope of the present discrete-element matrix theory only by the introduction of an imaginary box surrounding the atomic system under consideration, or by some similar awkward device. It is therefore of interest to examine an ingenious procedure suggested by Van Vleck[1] which permits the evaluation of second-order energy corrections for the lower energy levels without making the complete transformation V.

The essential feature of the method is the introduction of a different canonical transformation—we shall refer to it as a Van Vleck transformation—with a matrix $T(\lambda)$ which eliminates the terms of the first order in λ from the matrix components of H_1 connecting any particular unperturbed energy $E_k^{(0)}$ with the other unperturbed levels. The Van Vleck transformation is relatively easy to apply and makes it unnecessary to compute the difficult elements of H_1. The transformation depends on the particular unperturbed energy level $E_k^{(0)}$ to be dealt with, but for simplicity we suppress the index k in the notation for the matrix $T(\lambda)$.

Let us give $T(\lambda)$ the form[2]

$$T(\lambda) = e^{i\lambda S} = I + i\lambda S - \frac{1}{2}\lambda^2 S^2 - \frac{i\lambda^3 S^3}{6} + \cdots, \tag{48·22}$$

where S is Hermitian. The adjoint of T is

$$T(\lambda)^\dagger = e^{-i\lambda S} = I - i\lambda S - \frac{1}{2}\lambda^2 S^2 + \frac{i\lambda^3 S^3}{6} + \cdots. \tag{48·23}$$

As all powers of S are Hermitian, $TT^\dagger = T^\dagger T = I$, and T is automatically

---

[1] The method was first used in a paper by J. H. Van Vleck on "Sigma-type Doubling and Electron Spin," *Phys. Rev.* **33**, 467 (1929)—see especially pp. 484 and 485—but is more fully described by Jordahl, a pupil of Van Vleck, in a paper on paramagnetic susceptibility, *Phys. Rev.* **45**, 87 (1934).

[2] The particular form (48·22) was suggested to the author by Dr. Bela Lengyel.

unitary. Let $G(\lambda)$ denote the transformed Hamiltonian defined explicitly by

$$T^{-1}\tilde{H}T = G(\lambda) = G_0 + \lambda G_1 + \lambda^2 G_2 + \cdots. \qquad (48\cdot24)$$

Introducing (48·22), the equation for the determination of G takes the form

$$(I + i\lambda S - \tfrac{1}{2}\lambda^2 S^2 - \cdots)(G_0 + \lambda G_1 + \lambda^2 G_2 + \cdots)$$
$$= (H_0 + \lambda H_1 + \cdots)(I + i\lambda S - \tfrac{1}{2}\lambda^2 S^2 + \cdots). \qquad (48\cdot25)$$

Equating coefficients of like powers of λ

$$G_0 = H_0, \qquad\qquad\qquad\qquad\qquad\qquad\qquad (48\cdot26)$$
$$G_1 = H_1 + i(H_0 S - S G_0), \qquad\qquad\qquad (48\cdot27)$$
$$G_2 = H_2 + i(H_1 S - S G_1) - \tfrac{1}{2}(H_0 S^2 - S^2 G_0), \qquad (48\cdot28)$$

Let us next choose S so that $S(k''j;k'j') = 0$ if $k' \neq k$ and $k'' \neq k$. Evidently this choice does not affect the unitary character of T. It follows from (48·27) that under these circumstances $G_1(kj;kj')$ and $G_1(k''j'';k'j')$ are equal to $H_1(kj;kj')$ and $H_1(k''j'';k'j')$, respectively, provided neither k'' nor k' is equal to k. Thus the transformation T to terms of the second order in λ affects only that portion of H which connects $E_k^{(0)}$ with other unperturbed levels. We now introduce the requirement that elements of G_1 of the form $G_1(kj;k'j')$ with $k \neq k'$ shall all vanish. Equations (48·26) and (48·27) give

$$H_1(kj;k'j') + i[H_0(kj;kj) - H_0(k'j';k'j')]S(kj;k'j') = 0,$$

or

$$S(kj;k'j') = i\frac{H_1(kj;k'j')}{E_k^{(0)} - E_{k'}^{(0)}}. \qquad (48\cdot29)$$

It thus appears that our problem is solvable. Having determined S and G with the aid of (48·27) we can employ (48·28) and the equations of higher order to compute G_2, etc. In particular we find after reduction that

$$G_2(kj;kj') = H_2(kj;kj') + \sum_{k' \neq k}\sum_{j''}\frac{H_1(kj;k'j'')H_1(k'j'';kj')}{E_k^{(0)} - E_{k'}^{(0)}}. \qquad (48\cdot30)$$

In order to compute the second-order energy correction we now apply the method of Sec. 48b, which goes through more simply than before. First we apply a canonical transformation with matrix V which diagonalizes the minor $^{(k)}H_1$, i.e., $^{(k)}G_1$. Let $\bar{G}$ denote the transformed matrix $V^{-1}GV$. Equations (48·10) to (48·12) are now applicable with $\bar{G}_1$, and $\bar{G}_2$ substituted for H_1 and H_2, respectively. As before, we use $\bar{U}$ to denote the transformation matrix required to diagonalize $\bar{G}$. The

$(kl;k'i)$ elements of (48·11) show that $\widehat{U}^{(1)}(kl;k'i)$ is now zero when $k' \neq k$. In place of (48·17) we have

$$\widehat{U}^{(0)}(kl;ki)E_{ki}^{(2)} - \widehat{U}^{(1)}(kl;ki)(E_{kl}^{(1)} - E_{ki}^{(1)}) - \sum_{l''}\widehat{G}_2(kl;kl'')\widehat{U}^{(0)}(kl'';ki) = 0.$$

$$(48·31)$$

Choosing i and l to make $E_{ki}^{(1)}$ and $E_{kl}^{(1)}$ equal, we reduce the equations to the form of Eqs. (48·18) with $\widehat{G}_2(kl;kl'')$ substituted for $F(kl;kl'')$. The solution is to be carried through as before with the advantage that in this case the required elements of $\widehat{G}_2$ are obtained much more easily than the corresponding elements of $^{(ki)}F$. Whereas in the former case we had to apply the transformation V to elements of H_1 connecting $E_k^{(0)}$ with all other unperturbed levels, in this case we have no corresponding elements of G_1 to transform and, in applying V, we can deal with a single small square of G which belongs to the energy level $E_k^{(0)}$.

In case the higher order perturbations remove no degeneracy left by the first-order perturbation—i.e., in the case to which (48·21) is applicable—Eqs. (48·31) reduce to $E_{ki}^{(2)} = \widehat{G}_2(ki;ki)$.

48d. Simplification of Perturbation Calculations by Means of Integrals of the Perturbed Hamiltonian.—Fortunately the complexities of the general perturbation method can usually be reduced by proper use of the "integrals of the motion," i.e., the dynamical variables which commute with $\widetilde{H}$ (cf. Sec. 38c, p. 291).

Let us assume that for all values of λ, α and $\widetilde{H}$ form a normal commuting pair of dynamical variables (cf. Sec. 37c, p. 286) so that they have a complete system of simultaneous eigenfunctions. With these functions their matrices can be made simultaneously diagonal. We shall further suppose that a complete system of simultaneous eigenfunctions of α and the unperturbed Hamiltonian H_0 is given. Under these circumstances it is convenient to seek solutions of the perturbed problem which are simultaneous eigenfunctions of α and $\widetilde{H}$. We can then deal with one eigenvalue of α at a time as well as with one eigenvalue of H_0. Let $\widetilde{\psi}_a$ denote such a simultaneous eigenfunction of α and $\widetilde{H}$, the α eigenvalue being a. Then $\widetilde{\psi}_a$ is expressible as a linear combination of unperturbed functions for the same eigenvalue of α. All other unperturbed functions can be ignored in working out $\widetilde{\psi}_a$ and its $\widetilde{H}$ eigenvalue. Thus the original perturbation problem can be said to factor into a number of partial problems, one for each eigenvalue of α.

Let u_{knm} denote a simultaneous eigenfunction of H_0, and α with the respective eigenvalues $E_k^{(0)}$ and a_n. m is a degeneracy index indicating the eigenvalues of additional variables which together with H_0 and α make up a complete commuting set. Let H_1, H_2, $\cdots$ denote the matrices of the corresponding operators derived from the basic functions u_{knm}. The matrices of H_0 and α in this scheme are diagonal. We

denote them by $\mathbf{E}^{(0)}$ and $\mathbf{A}$, respectively. By hypothesis α commutes with $\tilde{H}$ for all values of λ. Hence it commutes with H_1, H_2, $\cdots$. Thus the diagonal matrix $\mathbf{A}$ commutes with $\mathbf{H}_1$, $\mathbf{H}_2$, $\cdots$, and consequently the latter matrices have no elements connecting different eigenvalues of α. In other words every term in the expansion of $\tilde{\mathbf{H}}$ in powers of λ is diagonal in the index n.

Let the matrix $\tilde{\mathbf{U}}$ convert $\tilde{\mathbf{H}}$ to diagonal form at the same time keeping the matrix of α diagonal. By proper numbering of rows and columns we can insure that the matrix of α is actually invariant of the canonical transformation with the matrix $\tilde{\mathbf{U}}$. But this means that $\tilde{\mathbf{U}}$ commutes with $\mathbf{A}$ and implies that $\tilde{\mathbf{U}}$, and each term in its expansion, is diagonal in n. In Eqs. (48·8), (48·9), etc., we can now replace each matrix by a submatrix obtained from the original one by picking out such elements as $H_1(k',n',m'; k'',n',m'')$ and holding n' fast while the other indices range through all possible values. The whole perturbation calculation can then be carried through separately for each value of n'. The advantage is, of course, that by this procedure we reduce the order of the secular determinants to be solved and in fact greatly reduce the number of matrix elements to be dealt with in solving the complete problem.

Evidently if there are several independent dynamical variables which commute with $\tilde{H}$ we can use simultaneous eigenfunctions of all of them and so deal with a single set of eigenvalues at a time. It is desirable, if possible, to pick out independent integrals of the motion α_1, α_2, α_3, $\cdots$ which, together with $\tilde{H}$, form a complete normal commuting set. Their simultaneous eigenfunctions will then be uniquely defined except for phase factors. If both the unperturbed and perturbed eigenfunctions are chosen to be simultaneous eigenfunctions of all these integrals, the secular determinants will be reduced to the lowest order possible. It is not necessarily possible to reduce the order of the secular determinants to unity, however, for the reason that the dynamical variables α_1, α_2, α_3, $\cdots$, which unite with $\tilde{H}$ to form a complete independent commuting set, may not, and generally do not, combine with H_0 to form a similar set. This is due to the fact that ordinarily H_0 has a higher symmetry than $\tilde{H}$. If such were not the case, the unperturbed problem would be no simpler than the perturbed one. Consequently every dynamical variable which commutes with H_0 need not necessarily commute with $\tilde{H}$. Moreover, dynamical variables which are independent of H_0 may not be independent of $\tilde{H}$, although they commute with both.

Let us assume that one or more additional independent dynamical variables β_1, β_2, $\cdots$ must be added to the α's in order to secure a complete independent commuting set H_0, α_1, $\cdots$, β_1, $\cdots$ for the unperturbed problem. Then either (a) all the β's commute with $\tilde{H}$ but are functions of $\tilde{H}$ and the α's, or (b) one or more of the β's fail to

commute with $\tilde{H}$. In the former case we can evidently deal with one set of eigenvalues of the α's and the β's at a time in applying the perturbation method—thus resolving the complete problem into a multiplicity of partial problems, each having nondegenerate sets of unperturbed eigenvalues. In this case the order of the secular determinants is reduced to unity and each partial problem can be dealt with by the methods of Sec. 47. On the other hand in case (b) each partial perturbation problem must necessarily involve a multiplicity of eigenvalues of those β's which do not commute with $\tilde{H}$. The zero-order wave functions are then unknown linear combinations of a multiplicity of initial unperturbed functions and the secular determinants cannot all be reduced to the first order. Both cases are illustrated in the examples of Secs. 49 and 50.

49. THE ENERGY LEVELS OF AN HYDROGENIC ATOM IN A UNIFORM MAGNETIC FIELD (SPIN NEGLECTED)[1]

49a. Derivation of Hamiltonian Operator.—As an elementary example of the operation of the perturbation method for an initially degenerate system let us consider the effect of a constant and uniform magnetic field on the energy levels and wave functions of an hydrogenic atom. Our conclusions would require revision if we were to take into account the electron spin, but the solution of the problem without the electron-spin terms is a useful exercise to carry through at this point. Let H_0 denote the unperturbed Hamiltonian

$$H_0 = -\frac{h^2}{8\pi^2\mu}\nabla^2 + \frac{Ze^2}{r}. \qquad (49\cdot1)$$

The complete Hamiltonian for an hydrogenic atom in an arbitrary external electromagnetic field with scalar potential Φ and vector potential $\vec{\alpha}$ is [cf. Eq. (7·11)]

$$H = H_0 + \frac{ehi}{2\pi\mu c}\left(\alpha_x\frac{\partial}{\partial x} + \alpha_y\frac{\partial}{\partial y} + \alpha_z\frac{\partial}{\partial z}\right) - \frac{ehi}{4\pi\mu c^2}\frac{\partial\Phi}{\partial t} + e\Phi + \frac{e^2}{2\mu c^2}|\vec{\alpha}|^2. \quad (49\cdot2)$$

We assume a uniform magnetic field of strength $\mathcal{K}$ in the direction of the positive z axis. Then

$$\text{curl } \vec{\alpha} = \vec{k}\mathcal{K}, \qquad (49\cdot3)$$

where $\vec{k}$ is a unit vector along the z axis. Since the physical results of our calculation will depend only on the curl of the vector potential[2] we may choose for $\vec{\alpha}$ the following simple particular solution of Eq. (49·3):

$$\alpha_x = -\tfrac{1}{2}\mathcal{K}y; \qquad \alpha_y = \tfrac{1}{2}\mathcal{K}x; \qquad \alpha_z = 0. \qquad (49\cdot4)$$

[1] Cf. L. BRILLOUIN, J. de Physique **8**, 74 (1927).
[2] Cf. discussion of "gauge invariance," p. 29.

Introducing spherical coordinates,

$$\alpha_x\frac{\partial}{\partial x} + \alpha_y\frac{\partial}{\partial y} + \alpha_z\frac{\partial}{\partial z} = \frac{\mathfrak{K}}{2}\left[x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right] = \frac{i\pi\mathfrak{K}\mathfrak{L}_z}{h} = \frac{\mathfrak{K}}{2}\frac{\partial}{\partial\varphi} \quad (49\cdot5)$$

and

$$|\vec{\alpha}|^2 = \frac{\mathfrak{K}^2(x^2 + y^2)}{4} = \frac{\mathfrak{K}^2}{4}r^2\sin^2\theta. \quad (49\cdot6)$$

Setting Φ equal to zero, the expression for H reduces to

$$H = H_0 - \frac{e\mathfrak{K}}{2\mu c}\frac{h}{2\pi i}\frac{\partial}{\partial\varphi} + \frac{e^2\mathfrak{K}^2}{8\mu c^2}r^2\sin^2\theta. \quad (49\cdot7)$$

We identify $-\dfrac{e\mathfrak{K}}{2\mu c}$ with the parameter λ of Secs. 47 and 48 and write H in the standard form

$$H = H_0 + \lambda H_1 + \lambda^2 H_2, \quad (49\cdot8)$$

with

$$H_1 = \mathfrak{L}_z = \frac{h}{2\pi i}\frac{\partial}{\partial\varphi}, \quad (49\cdot9)$$

$$H_2 = \mu r^2\sin^2\theta/2. \quad (49\cdot10)$$

For the magnetic fields obtainable experimentally the second-order term $\lambda^2 H_2$ is relatively small and unimportant spectroscopically. This term—for atoms and molecules in general—is essential, however, to the theory of diamagnetism.

49b. Legitimacy of the Perturbation Method.—The first question to be considered is the effect of the perturbative terms on the behavior of solutions of the Schrödinger equation at infinity. We know already that the unperturbed eigenfunctions having the standard form

$$u_{nlm} = \frac{R_{nl}\Theta_{lm}e^{im\varphi}}{\sqrt{2\pi}} \quad (49\cdot11)$$

are simultaneous eigenfunctions of H_0 and $\mathfrak{L}_z$. It follows that the addition of the operator $\mathfrak{L}_z$ or H_1 to the unperturbed Hamiltonian leaves the behavior of solutions at the singular points unchanged. In fact the functions u_{nlm} are exact eigenfunctions of the Hamiltonian $H_0 + \lambda H_1$.

H_2 increases the effective potential energy by an amount proportional to $r^2\sin^2\theta$. Thus the limit of the effective potential energy as r increases indefinitely is changed from zero to plus infinity in every direction except along the z axis. The form of the continuous-spectrum solutions will be radically altered in consequence and the discrete energy levels will be raised. There is nothing in this perturbative term to destroy the discrete spectrum, however, or to prevent the continuous transition of discrete eigenfunctions of $\tilde{H}$ into corresponding eigenfunctions of H_0 as λ approaches zero. We conclude that the application of the per-

turbation method to the determination of the discrete eigenfunctions in the presence of the field is legitimate.

49c. First-order Energy Correction; Relation to Magnetic Moment and Larmor Precession.—So long as H_2 is neglected, the problem of the influence of a uniform magnetic field on an hydrogenic atom is not one which requires perturbation methods. Since u_{nlm} is a simultaneous eigenfunction of H_0 and H_1 it is an eigenfunction of $H_0 + \lambda H_1$ with eigenvalue

$$E_{nlm} = E_n^{(0)} + \frac{\lambda mh}{2\pi} = -\frac{Nhc}{n^2} - \frac{meh\mathfrak{IC}}{4\pi\mu c}. \tag{49·12}$$

Thus, neglecting the secondary effects due to the square of the magnetic field strength, the result of the application of the field is to split each unperturbed energy level $E_n^{(0)}$ into $2n - 1$ equally spaced sublevels symmetrically placed with respect to $E_n^{(0)}$. In these formulas e denotes the algebraic value of the electronic charge, and the energy perturbation is therefore positive when m, or the component of angular momentum in the direction of the field, is positive.

From the expansions of Secs. 29d and 30i we know that the operators H_0, $\mathfrak{L}^2$, $\mathfrak{L}_z$ form a complete set of normally commuting independent operators. The unperturbed problem has Coulomb degeneracy superimposed on the normal degeneracy of the rotation-reflection group (*cf.* Secs. 40e and f). On the other hand every integral of the perturbed Hamiltonian $H_0 + H_1$ must commute with $\mathfrak{L}_z$ as well as with H_0. Rotations about axes other than the z axis are eliminated from the list of integrals of the Schrödinger equation. With them go the real observables $\mathfrak{L}_x$, $\mathfrak{L}_y$, R_x, R_y. As a result of a drastic reduction in the group of the Schrödinger equation every integral of the perturbed equation commutes with $\mathfrak{L}_z$, and $\mathfrak{L}_z$ becomes a function of H within the discrete spectrum. Thus the perturbation removes the degeneracy with respect to the different eigenvalues of $\mathfrak{L}_z$.

Comparing this concrete situation with the general discussion of Sec. 48d, p. 397, we note that the set of dynamical variables β_1, β_2, $\cdots$ need contain but one member, which can be identified with $\mathfrak{L}_z$, or with $\mathfrak{L}_x$, or $\mathfrak{L}_y$. If we identify β with $\mathfrak{L}_z$, we have an example of case (*a*) in which β commutes with $\bar{H}$ and the α's but is not independent of them. Otherwise we have an example of case (*b*). The same theory applied to an atomic model in which an electron moves in a central force field which does not obey the inverse-square law goes through as before except that each energy level has a definite value of $\mathfrak{L}^2$, so that there is no dynamical variable of the type of the α's of Sec. 48d.

The spacing of the sublevels into which $E_n^{(0)}$ is split by the field is $h\left|\dfrac{e\mathfrak{IC}}{4\pi\mu c}\right|$ according to (49·12). Here the cofactor of Planck's constant h is the frequency $\nu_{\mathfrak{IC}}$ of the classical Larmor precession in the field $\mathfrak{IC}$.

The classical vector magnetic moment $\overrightarrow{\mathfrak{M}}$ of one or more spinless electrons in orbital motion about a center of force with the vector angular momentum $\overrightarrow{\mathfrak{L}}$ is given by the formula

$$\overrightarrow{\mathfrak{M}} = \frac{e}{2\mu c}\overrightarrow{\mathfrak{L}} = -\frac{|e|}{2\mu c}\overrightarrow{\mathfrak{L}}. \qquad (49 \cdot 13)$$

If the electron is under the influence of a magnetic field $\mathfrak{R}$ in the direction of the z axis, the classical component of angular momentum along that axis is conserved and the mutual energy of the system and the field is $-\mathfrak{R}\mathfrak{M}_z$, or $-\dfrac{e\mathfrak{R}\mathfrak{L}_z}{2\mu c}$, in conformity with (49·12).

The existence of a quantum-mechanical Larmor precession of frequency

$$\nu_{\mathfrak{R}} = \frac{|e|\mathfrak{R}}{4\pi\mu c} \qquad (49 \cdot 14)$$

in the case of an hydrogenic atom with partially determinate azimuthal angle is readily derived from the above results. Since the value of φ for an assemblage of atoms with a unique value of $\mathfrak{L}_z$ is completely indeterminate, a wave function which partially specifies the instantaneous value of φ must involve a multiplicity of $\mathfrak{L}_z$ and energy values. It follows from Eqs. (49·12) and (49·14) that

$$\Psi = \sum_n \sum_l \sum_m c_{nlm} R_{nl} \Theta_{lm} e^{-\frac{2\pi i}{h} E_n^{(0)} t} e^{im(\varphi - \nu_{\mathfrak{R}} t)}. \qquad (49 \cdot 15)$$

Let us now compare the wave functions of two assemblages in the same state at $t = 0$, one of which moves under the influence of a magnetic field $\mathfrak{R}$, while the other is subject to no such field. If Eq. (49·15) gives the wave function of the former,

$$\Psi_1 = \sum_n \sum_l \sum_m c_{nlm} R_{nl} \Theta_{lm} e^{-\frac{2\pi i}{h} E_n^{(0)} t} e^{im\varphi} \qquad (49 \cdot 16)$$

describes the wave function of the latter. If we now refer the motion of the first assemblage to a set of coordinates r, θ, φ' which precesses about the z axis with the frequency $\nu_{\mathfrak{R}}$, the two wave functions take identical form. Thus the motion of the systems which are subject to the magnetic field, when referred to rotating coordinates, is exactly the same as the motion of those not subject to the field when referred to fixed coordinates. This is the exact quantum-mechanical equivalent of the classical Larmor theorem and can be used to derive the latter in the limiting case of a sharply defined wave packet.

In Sec. 55c it will be proved that the effect of a uniform magnetic field on the spectrum emitted by a hydrogenic-atom model with a spinless electron is also the same in quantum theory as in classical theory.

49d. The Second-order Energy Correction.—In determining the second-order energy correction we have to apply Eqs. (48·18), (48·19), and (48·20). As the initial matrix H_1 is diagonal, it is not necessary to apply the canonical transformation V which carries H_1 into H_1. Equation (48·19) reduces to

$$F(nlm;nl'm') = H_2(nlm;nl'm'). \tag{49·17}$$

When the perturbing term H_2 is taken into account, the Coulomb degeneracy is lifted and $\mathcal{L}^2$ ceases to be an integral of the motion. The only remaining integrals are $\mathcal{L}_z$, R_z, and functions of them. Since $\mathcal{L}_z$ and R_z commute, there is no degeneracy whatsoever in the discrete spectrum of the perturbed problem. Each perturbed level will have definite eigenvalues of $\mathcal{L}_z$ and R_z which can be used along with the initial total quantum number in labeling. The unperturbed wave functions u_{nlm} are symmetric or antisymmetric with respect to the operator R_z (which merely replaces θ by $\pi - \theta$) according as $l - m$ is even or odd. Thus the matrices of both these integrals of the perturbed problem are initially diagonal and we are in a position to apply the procedure of Sec. 48d. The matrix $\tilde{U}$ will have no elements connecting different eigenvalues of $\mathcal{L}_z$ and R_z and we can deal with one pair of eigenvalues of these operators at a time. (So far as the second-order correction is concerned the restriction to one eigenvalue of $\mathcal{L}_z$ at a time is already implied in Eqs. (48·19) and (48·20), since each different value of m gives a different first-order energy correction.) Every element $\tilde{U}(nlm; nl'm')$ vanishes unless $m' = m$ and $l - m$ and $l' - m$ are either both even, or both odd. Hence $l - l'$ must be an even number.

Let $^{(nm\tau)}H_2$ denote the minor of H_2 composed of elements of the type $H_2(nlm; nl'm)$ with fixed n and m and even or odd values of l and l' according as τ is set equal to $+1$, or -1. The second-order energy corrections are then to be computed from the secular equation

$$\det \left(^{(nm\tau)}H_2 - E^{(2)}I \right) = 0. \tag{49·18}$$

The required elements of H_2 are given by the formula

$$H_2(nlm;nl'm) = \frac{\mu}{4\pi} \int R_{nl} R_{nl'} \Theta_{lm} \Theta_{l'm} r^4 \sin^3\theta \; dr d\theta d\varphi. \tag{49·19}$$

Introducing the notation

$$\alpha(nl;nl') = \frac{\mu}{2} \int_0^\infty R_{nl} R_{nl'} r^4 dr,$$

$$\beta(lm;l'm) = \int_0^\pi \sin^3\theta \; \Theta_{lm}\Theta_{l'm} d\theta = \int_{-1}^{+1} (1 - x^2)\Theta_{lm}(x)\Theta_{l'm}(x) dx,$$

we reduce the above to

$$H_2(nlm;nl'm) = \alpha(nl;nl')\beta(lm;l'm). \tag{49·20}$$

Since $\Theta_{lm}(x)$ is independent of the sign of m,

$$H_2(nlm;nl'm) = H_2(nl - m; nl' - m).$$

The detailed calculation of these integrals would not be very difficult but is omitted in view of the fact that the energy corrections are so small that they have no great physical interest.[1] Suffice it to note that the diagonal elements of H_2 are all essentially positive; that for fixed values of l they increase rapidly with n and $|m|$; and that for fixed values of n and m they decrease as l increases. It follows from the invariance of the diagonal sum of the elements of a matrix with respect to a canonical transformation (*cf.* Sec. 44*b*) that the energy corrections $E^{(2)}$ are positive on the average and increase with n and $|m|$. Actually the energy corrections are individually positive, for each is the mean value of a positive function of the coordinates for the correct zero-order wave functions. The secular determinants met with in computing the second-order energy corrections for values of n up to and including $n = 4$ are never of order greater than 2.

The matrix $U^{(0)}$ is now derivable from (48·18) since in this case H_1 is diagonal and yields no information. From $U^{(0)}$ the zero-order wave functions are readily derived.

From the equation immediately preceding (48·18) and the diagonal character of H_1 we see that in our present problem all elements of $U^{(1)}$ connecting different unperturbed energies must vanish. Since $\tilde{U}$ has no nonvanishing components connecting different values of m, the matrix $U^{(1)}$ has no such components. The only components which can differ from zero are consequently those which are diagonal in both n and m. Equation (48·17) yields no information concerning them. In fact these elements are arbitrary except for the limitation imposed by (48·14). (This arbitrariness parallels that previously noted in connection with the nondegenerate case (*cf.* Sec. 47*a*, p. 383.) Equation (48·14) is most simply satisfied by setting the remaining elements of $U^{(1)}$ equal to zero. Thus we can identify the first-order wave functions with those of zero order.

50. THE ENERGY LEVELS OF AN HYDROGENIC ATOM IN A UNIFORM ELECTRIC FIELD

The problem of the splitting of the lines in the spectrum of an hydrogenic atom by a uniform external electric field (Stark effect) was first solved by Schwarzschild[2] and Epstein[3] in terms of the Bohr theory.

---

[1] *Cf.* O. HALPERN and TH. SEXL, *Ann. der Physik* **3**, 565 (1929); also J. H. VAN VLECK, *Electric and Magnetic Susceptibilities*, pp. 178–220, Oxford, 1932, for other references and for general discussion of the relation of these energy corrections to the problem of diamagnetism.

[2] K. SCHWARZSCHILD, *Berliner Sitzungsber.*, April, 1916, p. 548.

[3] P. S. EPSTEIN, *Ann. d. Physik* **50**, 489 (1916). The application of the correspondence principle to the determination of the relative intensities of the lines in the pattern is due to H. A. KRAMERS, *Danske Vidensk. Selsk. Skrifter* (8) **III**, **3**, 287, and *Zeits. f. Physik* **3**, 169 (1920).

The agreement of their results with the experimental Stark-effect patterns was one of the most important early successes of the Bohr theory. The same results to terms linear in the field strength were later derived by Schrödinger and by Epstein[1] from the standpoint of wave mechanics. Schrödinger solved the problem by separation of the variables in parabolic coordinates and also gave a partial discussion by the perturbation method without resorting to parabolic coordinates. As an illustration of the method of Sec. 48 we here follow Schrödinger's second procedure. The electron spin, which plays a secondary role for hydrogenic atoms in strong electric fields, is neglected. For a discussion which takes spin into account the reader is referred to the papers of Rojansky[2] and Schlapp.[3]

Let the field be of strength $\mathcal{E}$ and directed along the positive z axis. We can assume without loss of generality that the electrostatic potential Φ vanishes at the nucleus. The Hamiltonian of Eq. (49·2) reduces to

$$H = H_0 - e\mathcal{E}z = H_0 + \lambda H_1, \qquad (50·1)$$

with the conventions $\lambda = -e\mathcal{E}$; $H_1 = z$.

As in the problem of the external magnetic field all values of λ have experimental significance and we need not distinguish between the interpolation problem and the perturbed problem.

Implicit in Eq. (50·1) is the assumption that the electric field extends uniformly to infinity in all directions. Of course this is never true experimentally, but we know that the spectrum of an atom is empirically sensitive only to the field in its immediate neighborhood and the above assumption is simpler from an analytic point of view than any other. In spite of its simplicity, however, the assumption leads to a mathematical difficulty in the application of the perturbation method. The perturbative term in the Hamiltonian becomes negatively infinite if z approaches $+\infty$ when λ is positive, and if z approaches $-\infty$ when λ is negative. In the region of large negative values of the perturbing potential the Laplacian of a solution of the first Schrödinger equation is necessarily negative and it is therefore impossible for ψ to satisfy the boundary condition of quadratic integrability. In fact the perturbative term λz in the Hamiltonian has the effect of destroying the discrete spectrum of the problem and substituting for it a continuous spectrum of energy eigenvalues ranging from $-\infty$ to $+\infty$. It follows at once that the eigenfunctions and eigenvalues of the Hamiltonian H of Eq.

[1] E. Schrödinger, *Ann. d. Physik* **80**, 437 (1926); P. S. Epstein, *Phys. Rev.* **28**, 695 (1926).

[2] V. Rojansky, *Phys. Rev.* **33**, 1 (1929).

[3] R. Schlapp, *Proc. Roy. Soc.* **A119**, 313 (1928).

(50·1) do not continuously approach the eigenfunctions and eigenvalues of H_0 as λ approaches zero. Thus the fundamental postulate of the perturbation method (*cf.* Sec. 47*a*) is violated.

If the field is not too large, however, the situation bears a close resemblance to that of the Gamow-Gurney-Condon model for the alpha-particle disintegration of radioactive nuclei described in Sec. 31*b*. We have to do with a single particle under the influence of a potential V which is strongly negative in the neighborhood of the origin, and which either passes through a maximum as we move away from the origin along a radial line or else becomes infinite at infinity, depending on the direction. For not too large negative energies there are two regions of positive kinetic energy separated by a potential barrier. As one region is finite and the other extends to infinity, the conditions are right for weakly quantized energy levels.[1] The experimental spectrum shows that these levels exist and approach the discrete levels of H_0 as λ approaches zero. Spontaneous ionization of atoms by "leakage" through the potential barrier is a secondary phenomenon (*cf.* footnote 1, p. 181).

The wave functions of weakly quantized states are quadratically integrable solutions of the second Schrödinger equation which are almost, but not quite, monochromatic. They are accordingly approximate solutions of the first (time-free) Schrödinger equation. Hence, if we seek to solve the latter equation by successive approximations, using the perturbation method on a problem in which there are weakly quantized energy levels but no rigorously discrete energy eigenvalues, it is to be expected that early terms of the series will describe the imperfectly quantized states although the higher approximations do not converge in the normal manner. This expectation is confirmed by the result of applying perturbation theory to the problem in hand as well as by a study of the eigenvalue-eigenfunction problem by means of the variation method described in Sec. 51.

We proceed to the formal application of the procedure of Sec. 48. As in the magnetic case the group of the Schrödinger equation is reduced by the elimination of rotations about any but the z axis. $\mathcal{L}^2$ and R_z also drop out of the list of integrals, but R_x and R_y remain. Since neither R_x nor R_y commutes with $\mathcal{L}_z$, we infer that the perturbing energy does not completely remove the degeneracy of the unperturbed discrete energy levels. We use the same set of basic functions u_{nml} as in the discussion of the effect of a magnetic field in Sec. 49. There will then be no matrix components of H_1 involving different values of m and we can deal with one value of m at a time.

Since $z = r \cos \theta$, the general matrix element of H_1 is

$$H_1(nml;n'ml') = \int_0^\infty R_{nl}R_{n'l'}r^3dr\int_0^\pi \Theta_{lm}\Theta_{l'm} \cos \theta \sin \theta \, d\theta. \quad (50\cdot2)$$

[1] *Cf.* C. LANCZOS, *Zeits. f. Physik* **62**, 518 (1930), **65**, 431 (1930), **68**, 204 (1931).

The factor integral involving θ can be written in the form

$$\int_0^\pi \Theta_{lm}\Theta_{l'm} \cos\theta \sin\theta \, d\theta =$$

$$\left[\frac{(2l+1)(2l'+1)(l-m)!(l'-m)!}{4(l+m)!(l'+m)!}\right]^{1/2} \int_{-1}^{+1} P_{lm}(\xi)P_{l'm}(\xi)\xi d\xi, \quad (50\cdot3)$$

if we assume m to be positive. Negative values of m give the same matrix elements as the corresponding positive ones and it suffices to deal explicitly with the latter. The matrix elements diagonal in n are also readily seen to be independent of the sign of $l' - l$. For first-order calculations we need only the elements diagonal in n and can therefore restrict ourselves to an explicit discussion of the elements for which $l' - l$ is positive. By the recurrence formula (F-10) of Appendix F we have

$$\int_{-1}^{+1} P_{lm}(\xi)P_{l'm}(\xi)\xi d\xi = \frac{l+m}{2l+1}\int_{-1}^{+1} P_{l-1,m}(\xi)P_{l'm}(\xi)d\xi$$

$$+ \frac{l-m+1}{2l+1}\int_{-1}^{+1} P_{l+1,m}(\xi)P_{l'm}(\xi)d\xi. \quad (50\cdot4)$$

This is different from zero only if $l' - l = \pm 1$. Thus all matrix elements vanish except those for which $l' - l = \pm 1$. It will suffice to compute elements of the form $H_1(n,m,l; n,m,l-1)$.

Setting $l' = l - 1$ in (50·4) and applying (F-15), we obtain

$$\int_{-1}^{+1} P_{lm}(\xi)P_{l-1,m}(\xi)\xi d\xi = \frac{2(l+m)!}{(2l+1)(2l-1)(l-m-1)!}. \quad (50\cdot5)$$

Equation (50·3) now reduces to the form

$$\int_0^\pi \Theta_{lm}\Theta_{l-1,m} \cos\theta \sin\theta \, d\theta = \sqrt{\frac{(l-m)(l+m)}{(2l+1)(2l-1)}}. \quad (50\cdot6)$$

We turn next to the integral containing the radial part of the wave function. By (29·8) the normalized radial wave function is

$$R_{nl}(r) = \left[\frac{2Z}{na_0}\right]^{3/2} \sqrt{\frac{(n-l-1)!}{[(n+l)!]^3}} \frac{1}{\sqrt{2n}} x^l e^{-x/2} L_{n+l}^{2l+1}(x),$$

where $x = 2rZ/na_0$. Hence

$$\int_0^\infty R_{nl}(r)R_{n,l-1}(r)r^3dr =$$

$$\frac{a_0(n+l)(n-l-1)!}{4Z[(n+l)!]^3}\sqrt{n^2-l^2}\int_0^\infty e^{-x}x^{2l+2}L_{n+l}^{2l+1}(x)L_{n+l-1}^{2l-1}(x)dx. \quad (50\cdot7)$$

According to Eq. (G-13) of Appendix G the integral on the right has the value $-6n\frac{(n+l-1)!}{(n-l-1)!}[(n+l)!]^2$. Thus

$$\int_0^\infty R_{nl}(r)R_{n,l-1}(r)r^3 dr = -\frac{3}{2}\frac{na_0}{Z}\sqrt{n^2 - l^2}. \tag{50·8}$$

The complete matrix element is

$$H_1(n,m,l; n,m,l-1) = -\frac{3}{2}\frac{na_0}{Z}\sqrt{\frac{(l^2 - m^2)(n^2 - l^2)}{4l^2 - 1}}. \tag{50·9}$$

Let $^{(nm)}\mathbf{H}_1$ denote the minor of $\|H_1(nml; n'm'l')\|$ obtained by setting n' equal to n, and m' equal to m. The first-order energy corrections $E_{nm}^{(1)}$ for states with unperturbed energy $E_n^{(0)}$ and the quantum number m are the roots of the secular equation $\det(^{(nm)}\mathbf{H}_1 - E^{(1)}\mathbf{I}) = 0$. If the rows and columns of $^{(nm)}\mathbf{H}_1$ are arranged similarly according to increasing values of l and l', all elements vanish except those in the two lines bordering the principal diagonal. Thus the secular determinant has a simple form to which the term *continuant* has been applied. As an example we write down the explicit secular equation for $n = 5$ and $m = 2$:

$$\begin{vmatrix} E^{(1)} & C\sqrt{\dfrac{(3^2-2^2)(5^2-3^2)}{4\times 3^2-1}} & 0 \\[3mm] C\sqrt{\dfrac{(3^2-2^2)(5^2-3^2)}{4\times 3^2-1}} & E^{(1)} & C\sqrt{\dfrac{(4^2-2^2)(5^2-4^2)}{4\times 4^2-1}} \\[3mm] 0 & C\sqrt{\dfrac{(4^2-2^2)(5^2-4^2)}{4\times 4^2-1}} & E^{(1)} \end{vmatrix} = 0. \tag{50·10}$$

C has the value $\dfrac{3}{2}\dfrac{na_0}{Z} = \dfrac{3}{8}\dfrac{h^2}{\pi^2\mu e^2 Z}$. The general solution of these secular equations was first worked out by Schlapp[1] using a method describing in Appendix K. The roots of the general secular equation are

$$E^{(1)} = C(n - |m| - 1),\quad C(n - |m| - 3),\ \cdots,\quad -C(n - |m| - 1).$$

Thus $E^{(1)}$ is equal to C multiplied by an integer, say k, which ranges for fixed n and m from $-(n - |m| - 1)$ to $+(n - |m| - 1)$. However, different values of m give overlapping sets of energy values, so that, considering all values of m simultaneously, we can write

$$\lambda E_{nk}^{(1)} = \lambda Ck = -\frac{3}{8}\frac{h^2 \mathcal{E}}{\pi^2 \mu e Z}nk. \quad k = 0, \pm 1, \pm 2,\ \cdots,\ \pm(n - 1) \tag{50·11}$$

For each set of values of n, m, k there is just one wave function, but for given values of n and k the quantum number m takes on alternate integral values between the limits $\pm(n - |k| - 1)$. The multiplicity of the energy level $E_{nk}^{(0+1)} = E_n^{(0)} + E_{nk}^{(1)}$ is accordingly $n - |k|$.

Equation (50·11) is the well-known formula for the linear Stark effect in hydrogen confirmed for moderate fields by numerous experiments.

[1] *Loc. cit.*, footnote 3, p. 404.

For fields greater than 100 kilovolts per centimeter the quadratic and cubic terms also produce an appreciable effect. The reader is referred to the original papers[1] for a discussion of the theory of these higher order approximations.

51. THE VARIATIONAL METHOD

51a. Reduction of the Variational Problem to Algebraic Form.—An extremely valuable supplement to the power-series perturbation theory is to be found in the so-called *variational method*. This method is based on the reduction of eigenvalue-eigenfunction problems to variational form as demonstrated in Sec. 24. We there used the variational formulation of such problems as an aid in proving the completeness of the system of eigenfunctions of a large class of Sturm-Liouville problems. Here we employ it as a guide in the approximate numerical computation of discrete eigenvalues and their eigenfunctions.

It follows from the work of Sec. 32e that if H denotes a normal Hamiltonian operator, Hermitian with respect to class D, and if we define Q, N, J by the equations

$$Q[\psi] \equiv (H\psi,\psi), \qquad N[\psi] \equiv (\psi,\psi), \qquad (51\cdot1)$$

$$J[\psi,E] \equiv Q - EN, \qquad (51\cdot2)$$

the problem of finding the class D eigenfunctions of H and their eigenvalues is equivalent to that of solving the variational problem

$$\delta J[\psi,E] = 0, \qquad (51\cdot3)$$

subject to class D boundary-continuity conditions. Here E is a parameter which must be properly chosen in order that (51·3) shall have a solution. To every extremal ψ_m of Eq. (51·3) there corresponds an eigenvalue E_m of E such that $J[\psi_m,E_m] = 0$, or

$$E_m = \frac{Q[\psi_m]}{N[\psi_m]}. \qquad (51\cdot4)$$

Every such extremal ψ_m is a class D eigenfunction of H with the eigenvalue E_m, and conversely every class D eigenfunction of H is a solution of the variational problem. The extremals of J are also the extremals of Q/N.

We further assume the following propositions without rigorous proof.[2]

a. The lowest eigenvalue is an absolute minimum value of Q/N for class D comparison functions.

---

[1] Quadratic Stark effect: P. S. Epstein, *loc. cit.*; G. Wentzel, *Zeits. f. Physik* **38**, 527 (1926); I. Waller, *Zeits. f. Physik* **38**, 640 (1926); J. H. Van Vleck, *Proc. Nat. Acad. Sci.* **12**, 662 (1926). Cubic Stark effect: Theory worked out by Doi and published by Ishida and Hiyama, *Sci. Pap., Inst. Phys. Chem. Research, Tokyo* **9**, 1 (1928).

[2] Reasons for these assumptions will be found in Secs. 25a and 32i.

b. If a sequence of orthogonal eigenfunctions ψ_1, ψ_2, ψ_3, $\cdots$ with corresponding eigenvalues E_1, E_2, E_3, $\cdots$ is so numbered that $E_{k+1} \geqslant E_k$ for all values of k, and if all discrete eigenfunctions are contained in the manifold of linear combinations of the ψ's, it follows that $Q[\psi_k]/N[\psi_k]$ is the absolute minimum of Q/N for comparison functions ψ of class D, subject to the restriction that

$$(\psi,\psi_1) = (\psi,\psi_2) = \cdots = (\psi,\psi_{k-1}) = 0. \tag{51·5}$$

Let us reduce the variational problem to algebraic form by the introduction of a complete orthonormal system of functions φ_1, φ_2, φ_3, $\cdots$ all of which belong to the Hermitian domain of H. Let $\xi(k)$ denote the kth Fourier coefficient of a class D function ψ with respect to this system, *i.e.*, $\xi(k) = (\psi,\varphi_k)$. We seek to determine the vector $\vec{\xi} = (\xi(1), \xi(2), \cdots)$ so that ψ shall be an extremal of $J[\psi,E]$. Let $H(m,k)$ denote the matrix element $(H\varphi_k,\varphi_m)$. From the Hermitian character of H and the completeness relation for the φ's [*cf.* Eq. (22·32)] we can derive the Fourier coefficients of $H\psi$ with respect to the φ's. Thus

$$(H\psi,\varphi_k) = (\psi,H\varphi_k) = \sum_m \xi(m)H(m,k)^* = \sum_m H(k,m)\xi(m).$$

Using the completeness relation again, we derive

$$Q[\psi] = (H\psi,\psi) = \sum_k (H\psi,\varphi_k)(\psi,\varphi_k)^* = \sum_k \left[\sum_m \xi(k)^* H(k,m)\xi(m) \right]. \tag{51·6}$$

Furthermore

$$N[\psi] = \sum_k \sum_m \xi(k)^* \xi(m)\delta_{km} = \sum_k |\xi(k)|^2, \tag{51·7}$$

$$J[\psi,E] = Q[\psi] - EN[\psi] = \sum_k \sum_m \xi(k)^* \xi(m)[H(k,m) - E\delta_{km}]. \tag{51·8}$$

Thus the integrals Q, N, J are reduced to Hermitian quadratic forms in the parameters $\xi(1)$, $\xi(2)$. $\cdots$.

The first variation of ψ can now be written as $\delta\psi = \sum_{k=1}^{\infty} \varphi_k \delta\xi(k)$, where the $\delta\xi$'s are arbitrary complex numbers. The first variation of the integral J is obtained by differentiating the corresponding quadratic form with respect to the ξ's. We can throw it into the form

$$\delta J = \sum_m \delta\xi(m) \sum_k \xi(k)^*[H(k,m) - E\delta_{km}] + \sum_m \delta\xi(m)^* \sum_k \xi(k)[H(m,k) - E\delta_{mk}].$$

$$\tag{51·9}$$

If ψ is an extremal of J for arbitrary admissible variations $\delta\psi$, δJ must

vanish for arbitrary complex values of the $\delta\xi$'s. Hence it is necessary that the coefficient of every $\delta\xi(m)$ and every $\delta\xi(m)^*$ in (51·9) shall vanish, *i.e.*,

$$\sum_{k=1}^{\infty} \xi(k)[H(m,k) - E\delta_{mk}] = 0. \qquad m = 1, 2, 3, \cdots. \qquad (51\cdot10)$$

Conversely, it is easy to see (*cf.* p. 360) that a solution of Eqs. (51·10) for which the sums $\sum_k|\xi(k)|^2$, $\sum_k\left|\sum_m H(k,m)\xi(m)\right|^2$ are convergent determines a corresponding function $\psi(x)$ which belongs to the Hermitian domain of H, extremalizes $J[\psi,E]$, and is representable to an arbitrarily high degree of precision in the least-squares sense by a finite number of terms of the series $\sum_k\varphi_k\xi(k)$. The method would give discrete eigenfunctions outside class D if such existed. We assume, however, in accordance with the hypotheses of Sec. 32, that all discrete eigenvalues of H are equipped with class D eigenfunctions.

Except for a minor difference of notation these new equations are identical with (44·30). Thus the variation principle leads us directly back to the basic relations of the matrix theory.

51b. The Ritz Method.—The derivation of Eqs. (51·10) from the variation principle suggests the use of the Ritz method of successive approximations for their solution.[1] This is a direct method of dealing with problems in the calculus of variations without the use of Euler's differential equations.

Let the problem to be solved be that of finding extremals of $Q[\psi]$ subject to class D boundary-continuity conditions and to the requirement that the comparison functions shall be normalized to make $N[\psi] = 1$. The solutions will of course be normalized extremals of Q/N and J. Assuming that the first $k - 1$ eigenvalues and their eigenfunctions have been found, we formulate the problem of finding the kth eigenvalue E_k and its eigenfunction ψ_k as that of *minimizing* $Q[\psi]$ subject to the normalization condition and the orthogonality conditions (51·5).

The essential feature of the method is the construction of a sequence of normalized comparison functions $u^{(1)}$, $u^{(2)}$, $u^{(3)}$, $\cdots$ conforming to (51·5) or having the property that $\lim_{n\to\infty} (u^{(n)},\psi_r) = 0$ ($r = 1, 2, \cdots$, $k - 1$), and of such a character that for every finite value of n,

$$Q[u^{(n)}] \geqslant E_k,$$

whereas $\lim_{n\to\infty} Q[u^{(n)}] = E_k$. Such a succession of functions is called a

---

[1] W. RITZ, *J. f. reine u. angew. Math.* **135**, 1–61 (1909); COURANT-HILBERT, *M.M.P.*, p. 150.

minimal sequence. When such a sequence is constructed one looks to the function $u^{(\infty)} = \lim\limits_{n \to \infty} u^{(n)}$ for a solution of the problem. The method is not necessarily successful in every case, but if the limit function $u^{(\infty)}$ exists and is an admissible comparison function, we can ordinarily expect that it will be a genuine extremal. Clearly the chance that the minimal sequence will converge on a definite limit function will be greatly enhanced if we can formulate the minimum problem in such a fashion as to insure that it has a unique solution, *i.e.*, that the eigenvalue E_k shall not have a multiplicity of linearly independent eigenfunctions which belong to the class of admissible comparison functions for the minimum problem. Ordinarily the desired uniqueness can be artificially created by imposing suitable symmetry conditions on the comparison functions of each of the minimum problems.

Our procedure is based on postulates a and b of pp. 408 and 409. Let us first consider the lowest energy level E_1. The orthogonality conditions (51·5) then drop out of the problem.

In order to construct the desired minimal sequence we make use of the orthonormal set of φ's introduced in Sec. 51a. Let $w^{(n)}$ denote the *arbitrary* n-term linear combination

$$w^{(n)}(x) = \sum_{k=1}^{n} \varphi_k(x)\xi(k). \tag{51·11}$$

Let $v_1^{(n)}$ denote the special n-term linear combination

$$v_1^{(n)}(x) = \sum_{k=1}^{n} \varphi_k(x)\eta_1(k), \tag{51·12}$$

for which $\eta_1(k) = (\psi_1, \varphi_k)$. The sequence

$$\frac{v_1^{(1)}(x)}{\sqrt{N[v_1^{(1)}]}}, \frac{v_1^{(2)}(x)}{\sqrt{N[v_1^{(2)}]}}, \frac{v_1^{(3)}(x)}{\sqrt{N[v_1^{(3)}]}} \cdots$$

is easily proved to be a minimal sequence for E_1. In order to establish this fact we make use of (51·6), identifying ψ with ψ_1 and replacing $\xi(k)$ by $\eta_1(k)$ throughout. The scalar products $\sum\limits_{m} H(k,m)\xi(m)$ and $\sum\limits_{k} (H\psi,\varphi_k)(\psi,\varphi_k)^*$ are absolutely convergent. It follows that for every positive ϵ, however small, there exists a value of n such that

$$\left| Q[\psi_1] - \sum_{k=1}^{n} \sum_{m=1}^{n} H(k,m)\eta_1(m)\eta_1(k)^* \right| < \epsilon,$$

and

$$\left| N[\psi_1] - \sum_{k=1}^{n} \eta_1(k)\eta_1(k)^* \right| < \epsilon.$$

Hence

$$\frac{Q[\psi_1]}{N[\psi_1]} = Q[\psi_1] = \frac{\sum_{k=1}^{n} \sum_{m=1}^{n} H(k,m)\eta_1(m)\eta_1(k)^*}{\sum_{k=1}^{n} \eta_1(k)\eta_1(k)^*} + O(\epsilon),$$

where $O(\epsilon)$ is a quantity of the order of magnitude of ϵ. But it follows from (51·6) that

$$\frac{Q[v_1^{(n)}]}{N[v_1^{(n)}]} = \frac{\sum_{k=1}^{n} \sum_{m=1}^{n} H(k,m)\eta_1(m)\eta_1(k)^*}{\sum_{k=1}^{n} \eta_1(k)\eta_1(k)^*} + O(\epsilon).$$

Therefore $\lim_{n \to \infty} \dfrac{Q[v_1^{(n)}]}{N[v_1^{(n)}]} = Q[\psi_1] = E_1$. We conclude that the functions $v_1^{(n)}/\sqrt{N[v_1^{(n)}]}$ fulfill the requirements for a normalized minimal sequence. (Because E_1 is the absolute minimum of $Q[\psi]$ the requirement $Q\left[\dfrac{v_1^{(n)}}{\sqrt{N[v_1^{(n)}]}}\right] \geqslant E_1$ is automatically satisfied.)

However, since we do not know ψ_1, we cannot actually set up this particular sequence of functions. Instead we introduce a different sequence, whose typical member is

$$u_1^{(n)}(x) = \sum_{k=}^{n} \varphi_k(x)\xi_1^{(n)}(k),$$

the coefficients $\xi_1^{(n)}(k)$ being obtained by minimizing $Q[w^{(n)}(x)]$ with respect to the parameters $\xi(k)$ subject to the normalization condition

$$N[w^{(n)}(x)] = \sum_{k=1}^{n} |\xi(k)|^2 = 1.$$

If $u_1^{(u)}(x)$ is constructed in this manner, it follows that

$$Q[u_1^{(n)}] \leqslant Q\left[\frac{v_1^{(n)}}{\sqrt{N[v_1^{(n)}]}}\right].$$

Consequently the $u_1^{(n)}$'s must also form a normalized minimal sequence for E_1.

In order to find the coefficients $\xi_1^{(n)}(k)$ we make use of the relation

$$Q[w^{(n)}(x)] = \sum_{k=1}^{n} \sum_{m=1}^{n} H(k,m)\xi(m)\xi(k)^*. \tag{51·13}$$

To derive the minimum value of this expression subject to the normalization condition

is equivalent to finding a normalized vector $\vec{\xi}_1^{(n)}$ which minimizes

$$J[w^{(n)},E] = \sum_{k=1}^{n} \sum_{m=1}^{n} [H(k,m) - E\delta_{mk}]\xi(m)\xi(k)^{*}.$$

Setting the derivatives of $J[w^{(n)},E]$ with respect to the real and imaginary parts of each coefficient $\xi(k)$ equal to zero, one obtains the finite set of n equations in n unknowns,

$$\sum_{m=1}^{n} \xi(m)[H(k,m) - E\delta_{mk}] = 0. \qquad k = 1, 2, \cdots n \qquad (51\cdot14)$$

Equations (51·10) form an obvious extrapolation of Eqs. (51·14) to the case where n is infinite.

Equations (51·14) have a nontrivial solution if, and only if, E is a root of the secular equation

$$\det\|H(k,m) - E\delta_{mk}\| = 0. \qquad (51\cdot15)$$

The extreme values of J are all zero (*cf.* Appendix E). Let $E_\lambda^{(n)}$ be a root of (51·15) and let $\vec{\xi}_\lambda^{(n)}$ be the corresponding normalized solution of (51·14). Finally, let

$$u_\lambda^{(n)} = \sum_{k=1}^{n} \varphi_k(x)\xi_\lambda^{(n)}(k).$$

Then $Q[u_\lambda^{(n)}]/N[u_\lambda^{(n)}] = Q[u_\lambda^{(n)}] = E_\lambda^{(n)}$. The minimum value of $Q[w^{(n)}]/N[w^{(n)}]$ must therefore be the lowest of the n roots of Eq. (51·15), which we call $E_1^{(n)}$. The corresponding set of functions $u_1^{(n)}$ form the desired minimal sequence, and it follows that $\lim\limits_{n\to\infty} E_1^{(n)} = E_1$.

We have found that the lowest eigenvalue of H is the limit as n approaches infinity of the lowest root of the nth-order secular equation (51·15). As the order n increases, the root $E_1^{(n)}$ approaches E_1 monotonically from above. This follows from the fact that the linear manifold M_n, defined by the first n φ's, includes M_{n-1} with the result that the minimum value of Q for comparison functions in M_n is less than or equal to the minimum value for comparison functions of M_{n-1}. Thus

$$E_1^{(n-1)} > E_1^{(n)} > E_1^{(n+1)} > \cdots > E_1.$$

If the eigenvalue E_1 is nondegenerate, the minimal sequence $u_1^{(1)}$, $u_1^{(2)}$, $u_1^{(3)} \cdots$ can be so chosen that it converges in the mean-square sense upon an eigenfunction of E_1. To prove this statement, let us expand $u_1^{(n)}$ in terms of a complete orthonormal set of eigenfunctions of H, say $\psi_1, \psi_2, \cdots$. (In order to make the formulas more transparent we write them in the form appropriate to the case where H has a purely discrete spectrum.) Thus, if $c_k^{(n)}$ is the Fourier coefficient $(u_1^{(n)},\psi_k)$, we have

$$(Hu_1^{(n)}, \psi_k) = (u_1^{(n)}, H\psi_k) = E_k(u_1^{(n)}, \psi_k) = E_k c_k^{(n)},$$

$$Q[u_1^{(n)}] = (Hu_1^{(n)}, u_1^{(n)}) = \sum_k E_k |c_k^{(n)}|^2 = E_1 \left[|c_1^{(n)}|^2 + \sum_{k>1} \left(\frac{E_k}{E_1} \right) |c_1^{(n)}|^2 \right].$$

Since $\sum_{k=1}^{\infty} |c_k^{(n)}|^2 = 1$, we have

$$E_1 = \lim_{n \to \infty} Q[u_1^{(n)}] = E_1 \left\{ 1 + \lim_{n \to \infty} \left[\sum_{k=2}^{\infty} \left(\frac{E_k}{E_1} - 1 \right) |c_k^{(n)}|^2 \right] \right\}$$

or

$$\lim_{n \to \infty} \left[\sum_{k=2}^{\infty} \left(\frac{E_k}{E_1} - 1 \right) |c_k^{(n)}|^2 \right] = 0.$$

But since the lowest energy level is nondegenerate,

$$\left(\frac{E_k}{E_1} - 1 \right) \geqslant \frac{E_2}{E_1} - 1 > 0.$$

Hence $\lim_{n \to \infty} \left[\sum_{k=2}^{\infty} |c_k^{(n)}|^2 \right] = 0$, and $\lim_{n \to \infty} |c_1^{(n)}|^2 = 1$. Let us next assume that ψ_1 is real and that in some portion of configuration space, say G, it is positive. We can choose $u_1^{(n)}$ also to be real and for sufficiently large n we can make $u_1^{(n)}$ positive in G. Under these circumstances it is clear that $\lim_{n \to \infty} c_1^{(n)} = 1$, so that the Fourier coefficients of $u_1^{(n)}$ converge on those of ψ_1.

The norm of the function $\psi_1 - u_1^{(n)}$ is now seen to converge upon the limit zero as n becomes infinite. Thus

$$\lim_{n \to \infty} (\psi_1 - u_1^{(n)}, \psi_1 - u_1^{(n)}) = \lim_{n \to \infty} \left[\sum_{k=2}^{\infty} |c_k^{(n)}|^2 + |c_1^{(n)} - 1|^2 \right]$$

$$= \lim_{n \to \infty} [|c_1^{(n)}|^2 - c_1^{(n)} - c_1^{(n)*} + 1] = 0.$$

This proves that the sequence of functions $u_1^{(n)}(x)$ does actually converge in the root-mean-square sense on the eigenfunction ψ_1. For quantum-mechanical purposes this is all that we require.

There is nothing in the above work which forbids us to impose additional restrictions on the comparison functions provided that the Hamiltonian H has eigenfunctions which satisfy these restrictions. Let us assume that H, together with the integrals α_1, α_2, $\cdots$, forms a complete set of commuting independent dynamical variables (*cf.* Sec. 37d). Then there will never be more than one simultaneous eigenfunction of H and the α's with a given set of eigenvalues. If we can find a complete

system of functions φ_1, φ_2, $\cdots$ which are simultaneous eigenfunctions of the α's, it is easy to deal with one set of eigenvalues of the α's at a time by imposing on the comparison functions of the minimum problem from the beginning the requirement that they shall be eigenfunctions of the α's with chosen eigenvalues. We then proceed as before using only those φ's which conform to the new restrictions. In this way it is possible to compute to any desired approximation the lowest eigenvalue of H having given symmetry properties. The corresponding eigenfunction is sure to be uniquely defined, except for a phase factor, so that a definite limit function can be attained.

*51c. Higher Roots of the Secular Equation.—The question now arises whether we cannot derive the higher eigenvalues and their eigenfunctions by a practical modification of the same procedure. The answer is that we have only to use the higher roots of the secular equation (51·15).[1] The kth root of that equation when the roots are arranged in ascending order of magnitude (f-fold roots appearing f times) approaches the corresponding eigenvalue of H as a limit when n goes to infinity, and does so monotonically from above as in the preceding case.

It will be sufficient to consider the evaluation of the second eigenvalue E_2 with its eigenfunction ψ_2. Let $\epsilon_2^{(n)}$ denote the minimum value of $Q(\psi)$ for all normalized comparison functions of the manifold M_n which are orthogonal to ψ_1—as required by Eq. (51·5). Clearly the sequence of values of $\epsilon_2^{(n)}$ approaches E_2 monotonically from above when $n \to \infty$ —we can use the same arguments as in the preceding discussion of $E_1^{(n)}$. The corresponding root of the secular equation (51·15), *viz.*, $E_2^{(n)}$, is easily shown to be the minimum value of $Q[\psi]$ for all normalized comparison functions which belong to M_n and are at the same time orthogonal to $u_1^{(n)}$.[2] We assume that $E_1^{(n)}$ is a simple root of the secular equation for all values of n and that in consequence $u_1^{(n)}$ approaches ψ_1 as a limit when n becomes infinite. (Difficulties arising in the multiple-root case are believed to be purely formal.) Hence

$$\lim_{n \to \infty} (u_2^{(n)}, \psi_1) = \lim_{n \to \infty} (u_2^{(n)}, u_1^{(n)}) = 0.$$

Thus the orthogonality condition used for $E_2^{(n)}$ approaches that for $\epsilon_2^{(n)}$ as n becomes very large and we can infer that $\lim_{n \to \infty} E_2^{(n)} = \lim_{n \to \infty} \epsilon_2^{(n)} = E_2$. It is a known theorem[3]

[1] *Cf.* E. A. HYLLERAAS and B. UNDHEIM, *Zeits. f. Physik* **65**, 759 (1930).

[2] *Cf.* COURANT-HILBERT, *M.M.P.*, Kap I, §3, pp. 20–24. It will be observed that the eigenvectors $\xi_\lambda^{(n)}$ can be used to form the columns of a unitary matrix U which transforms the $n \times n$ element matrix

$$\mathbf{H}^{(n)} \equiv \begin{Vmatrix} H(1,1) & \cdots & H(1,n) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ H(n,1) & \cdots & H(n,n) \end{Vmatrix}$$

to diagonal form and reduces $Q[u_\lambda^{(n)}]$ to a sum of squares. The eigenvalues $E_\lambda^{(n)}$ are the corresponding diagonal elements of the transformed matrix.

[3] *Cf.* COURANT-HILBERT, *M.M.P.*, Kap. I, §4, 1, pp. 26–29.

that $E_2^{(n)}$ is the maximum of all minima obtained by minimizing $Q[w^{(n)}]$ subject to
$$\sum_{m=1}^{n} |\xi(m)|^2 = 1$$ and any orthogonality condition of the form $(\vec{\xi}, \vec{f_i}) = 0$. It follows
that $E_2^{(n)}$ is always greater than or equal to $\epsilon_2^{(n)}$. Hylleraas and Undheim[1] have
proved that $E_2^{(n)}$ (more generally, $E_k^{(n)}$) approaches its limit monotonically from
above. Similarly it is easy to show that each of the roots of the secular equation
(51·15) yields a minimal sequence $u_k^{(1)}, u_k^{(2)}, u_k^{(3)}, \cdots$. In general each of these
sequences yields an eigenvalue and eigenfunction. But if the number of discrete ortho-
gonal eigenfunctions of the operator H is finite, say p, we may expect that the minimal
sequence for a root $E_k^{(n)}$ for which $k > p$ will not converge on any limit function.

**51d. General Observations Regarding the Use of the Variational
Method.**—In practice the ultimate rigorous convergence of the successive
Ritz approximations for E_k and ψ_k is of no great importance. General
formulas for $E_k^{(N)}$ and $u_k^{(n)}$ are not ordinarily obtainable and the work of
evaluating $E_k^{(n)}$ and $u_k^{(n)}$ increases rapidly with n. Hence the theoretical
physicist must ordinarily content himself with the numerical computation
of approximations for relatively small values of n. The success of such
calculations, like those made by the perturbation method, rests on a
happy choice of the basic sequence of functions $\varphi_1, \varphi_2, \varphi_3, \cdots$.

The fact that $E_k^{(n)}$ approaches E_k monotonically from above is of the
greatest importance in the location of computational errors and especially
in the comparison of computed and experimental energy values. If a
value computed by the above outlined variational method lies below the
experimental one, it follows that there has been a mistake in the com-
putation, or that the calculation was made on an inadequate, or incorrect,
basis.

It is evident that the convergence of the sequence of Ritz approxima-
tions will be most rapid when the basic functions $\varphi_1, \varphi_2, \varphi_3, \cdots$ are
themselves approximate eigenfunctions of H, perhaps obtained by the
exact solution of the eigenvalue-eigenfunction problem for a simplified
operator H_0 not very different from H. In this case we can employ
either the Ritz method or the perturbation method and a comparison
of the two schemes is illuminating. Consider, for example, a case in
which the operator H_0 has degenerate energy levels (*cf.* Sec. 48). Inspec-
tion of the perturbation scheme for computing the first-order energy
corrections and the zero-order wave functions shows that it is equivalent
to the following. Let the unperturbed level $E_k^{(0)}$ have g-fold degeneracy
and let $^{(k)}\mathbf{H}$ and $^{(k)}\mathbf{H}_0$ denote the $g \times g$ element matrices obtained by
picking out those elements of $\mathbf{H}$ and $\mathbf{H}_0$ which involve only the unper-
turbed wave functions u_{kj} of the level $E_k^{(0)}$. Thus

$$^{(k)}H(m,n) = \int u_{km}^* H u_{kn} d\tau. \qquad m,n = 1, 2, \cdots, g \qquad (51·16)$$

[1] *Loc. cit.*, footnote 1, p. 415.

Let $^{(k)}H_1 = {^{(k)}H} - {^{(k)}H_0}$. The first-order energy corrections $E_{ki}^{(1)}$ are the eigenvalues of $^{(k)}H_1$ and the total perturbed energies to first-order corrections (setting the parameter λ of Sec. 48 equal to unity) are the eigenvalues of $^{(k)}H$. The zero-order wave functions are linear combinations of the g functions $u_{k1}, u_{k2}, \cdots, u_{kg}$ derivable from the normalized eigenvectors of $^{(k)}H$ or $^{(k)}H_1$. These two matrices have the same eigenvectors since their difference $^{(k)}H_0$ is a multiple of the unit matrix I.

Thus in this approximation the perturbation theory gives the same eigenvalues and eigenfunctions as would be obtained by applying the variational method to the location of the extremals of $Q[\psi]$ for normalized functions of the form

$$\psi = \sum_{j=1}^{g} u_{kj}\xi_j.$$

It will be observed, however, that in *such* an application of the variational method the mth root of the secular equation does not give a good approximation of the mth eigenvalue of H. On the contrary it gives an approximation of the mth sublevel of the parent unperturbed level E_k^0.

We may conclude that, although—in the notation of p. 415—

$$\lim_{n\to\infty} E_k^{(n)} = E_k, \qquad \lim_{n\to\infty} u_k^{(n)} = \psi_k,$$

if we break off computations with some finite value of n the final approximation $u_k^{(n)}$ may resemble $\psi_{k+\tau}(\tau > 0)$ much more closely than ψ_k. Under these circumstances $E_k^{(n)}$ will approximate $E_{k+\tau}$ rather than E_k. Thus the interpretation of the results obtained when we compute a Ritz approximation with a finite number of basic functions $\varphi_1, \varphi_2, \cdots, \varphi_n$ depends essentially on our choice of these functions. If any of the φ's, say φ_k, approximates from the beginning one of the eigenfunctions of H, say ψ_l, we may expect that one of the $u^{(n)}$'s will also approximate ψ_l. The corresponding eigenvalue of $H^{(n)}$ will then approximate E_l. It will be evident to the reader that, powerful as this method is, it must be used with judgment and discrimination. Fortunately the experimental energies are usually available as a guide.

In the event that the basic set of φ's consists of the eigenfunctions of an unperturbed operator H_0 it is important to know *a priori* just which members of the set are most likely to contribute largely to any particular exact eigenfunction of the perturbed operator H. A partial answer to this question is given by the standard perturbation theory of Secs. 47 and 48 which shows that, if the perturbed energy value E_{ki} originates in the unperturbed eigenvalue $E_k^{(0)}$, the most important terms in the expansion

$$\tilde{\psi}_{ki} = \sum_n \sum_j u_{nj} \tilde{U}(nj;ki) \tag{51·17}$$

for small values of λ are, (a) those involving the unperturbed wave functions u_{kj} of $E_k^{(0)}$ and, (b) those involving functions u_{nj} for which $|E_n^{(0)} - E_k^{(0)}|$ is small. This statement ignores the differences in the magnitude of the different components of $\mathbf{H}$, but to take these into account usually accentuates somewhat the relative importance of the small values of $|E_n^{(0)} - E_k^{(0)}|$. Thus, in general, $\tilde{H}(nj;km)$ is small when $|E_k^{(0)} - E_n^{(0)}|$ is large, and, according to Eq. (48·19), this effect works in such a sense as to reduce the importance of terms which are already small because of the factor $(E_k^{(0)} - E_n^{(0)})$ in the denominator. The reader should bear in mind, however, that although the individual contributions to the energy of terms with large $|E_k^{(0)} - E_n^{(0)}|$ is apt to be small, the number of such terms is very large, so that their importance in the aggregate may be considerable.

51e. Modifications of the Method; Construction of Eigenfunctions from Non-orthogonal System.—Hitherto we have used the variational method only as a means of finding the best approximate eigenfunctions of H which are linear combinations of a finite number of terms of a basic normal orthogonal set. The requirement that the basic functions shall be normalized and mutually orthogonal is not essential, however, and in many cases proves inconvenient. If this requirement is dropped and the argument is carried through as before, one finds that $Q[w^{(n)}]$ is given by (51·13) as before, but that

$$N[w^{(n)}] = \sum_{m=1}^{n} \sum_{k=1}^{n} \xi(m)^*\xi(k)\sigma(m,k),$$

where $\sigma(m,k) = (\varphi_k, \varphi_m)$. Instead of Eqs. (51·14) and (51·15) one obtains

$$\sum_{k=1}^{n} \xi(k)[H(m,k) - E\,\sigma(m,k)] = 0, \qquad m = 1, 2, 3, \cdots, n \quad (51·18)$$

and

$$\det\|H(m,k) - E\,\sigma(m,k)\| = 0, \qquad (51·19)$$

respectively. These equations are not quite so easy to work with, but the degree of the secular equation in the unknown E is unaltered.

A still wider extension of the fundamental idea of the variational method is often used. Instead of finding the extreme values of $Q[\psi]$ for a class of comparison functions that contain certain parameters—the ξ's—linearly, we can introduce a flexible ψ function involving nonlinear parameters $\lambda_1, \lambda_2, \cdots, \lambda_n$, evaluate Q as a function of these parameters, and determine its minimum or extreme values. This procedure leads to a set of nonlinear equations for the λ's, viz.,

$$\frac{\partial}{\partial\lambda_k}Q(\lambda_1, \cdots, \lambda_n) = 0, \qquad k = 1, 2, \cdots, n$$

which must ordinarily be solved by cut-and-try methods.

Usually the number of nonlinear parameters which can be introduced to advantage is small and hence it is a common procedure to evaluate Q only for a few explicit sets of λ values, resorting to interpolation in order to locate the minimum. In the absence of a convenient scheme for orthogonalizing the comparison functions for the relative minimum problem associated with an upper energy level with respect to the eigenfunctions of lower levels, the use of nonlinear parameters is ordinarily restricted to the location of the lowest energy eigenvalue of a given symmetry type.

52. THE PROBLEM OF THE HYDROGEN MOLECULE

52a. The Fixed-nuclei Problem.—As an example of the application of the variational method to the approximate evaluation of energy levels, let us consider the simplest of molecular problems, *viz.*, that of the hydrogen molecule.

Owing to the large ratio of the nuclear mass to the electronic mass it is possible in first approximation to resolve the eigenvalue-eigenfunction problem for a diatomic molecule into two parts.[1] In the first stage of the computation one treats the nuclei as fixed centers of force and evaluates the eigenvalues $E^{(e)}$ and eigenfunctions $\psi^{(e)}$ of the simplified Hamiltonian for the electronic motion in the presence of the stationary nuclei. The energies derived in this way are functions of the internuclear distance r which also appears as a parameter in the electronic wave function $\psi^{(e)}$. In the second stage of the computation one evaluates the energies $E^{(n)}$ and wave functions $\psi^{(n)}$ of a two-particle nuclear problem in which $E^{(e)}(r)$ plays the part of potential energy[2] and a term is added for the gyroscopic effect of the electronic angular momentum if necessary. Each electronic state, *i.e.*, each $\psi^{(e)}$, when properly defined, yields its own nuclear problem and a corresponding complete system of nuclear wave functions. If any particular electronic state A has no component of angular momentum along the internuclear axis, the corresponding nuclear problem is non-gyroscopic and reduces to the problem of the "dumbbell" model discussed in Secs. 28h and 47c. An approximate complete eigenfunction of any definite state of electronic and nuclear motion is of the form $\psi^{(e)}\psi^{(n)}$.

We shall interest ourselves here in the first stage of the molecular problem, *viz.*, with the evaluation of the energies and wave functions for

---

[1] For an elementary classical discussion of this resolution of the molecular problem, *cf.* E. C. Kemble, "Report on Molecular Spectra in Gases," *Nat. Research Council,* 1926, p. 293. The quantum-mechanical justification was first given by Born and Oppenheimer, *Ann. d. Physik* **84,** 457 (1927). *Cf.* also R. de L. Kronig, *Zeits. f. Physik* **50,** 347 (1928); J. H. Van Vleck, *Phys. Rev.* **33,** 467 (1929).

[2] We assume that the Coulomb potential of the bare nuclei is included in the Hamiltonian of the first-stage electronic problem.

fixed nuclei.   In this stage lies the problem of the chemical union of the atoms.   If.such union is possible for any given electronic state, the function $E^{(e)}(r)$ should have the form of the curve for $V(r)$ shown in Fig. 10, and the dissociation energy D, aside from a small correction for the minimum energy of nuclear vibration, should be equal to the difference $E^{(e)}(\infty) - E^{(e)}(r_0)$.   The moment of inertia of the molecule in a vibrationless state and its classical frequency of vibration are also derivable from the graph of $E^{(e)}(r)$ when once that function has been evaluated.
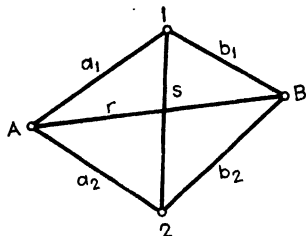
FIG. 21.—Coordinates for the hydrogen-molecule problem.

In the case of molecular hydrogen we have two nuclei which we designate as A and B and two electrons which we designate by the numbers 1 and 2 (cf. Fig. 21).   Let a_1 and b_1 denote the distances of electron 1 from the nuclei A and B, respectively.   Let a_2 and b_2 denote the corresponding distances for electron 2.   Let r and s denote the internuclear and interelectronic distances, respectively.   The basic differential equation for the fixed-nuclei problem is conveniently expressed in terms of Hartree's atomic units[1] in which it takes the form

$$H\psi \equiv -\frac{1}{2}(\nabla_1{}^2 + \nabla_2{}^2)\psi + \left(\frac{1}{r} + \frac{1}{s} - \frac{1}{a_1} - \frac{1}{a_2} - \frac{1}{b_1} - \frac{1}{b_2}\right)\psi = E^{(e)}\psi,$$

(52·1)

if we omit the electron-spin terms as usual.   In the future we shall discard the superscript $^{(e)}$ and designate the energy in this problem by E.

52b. The Heitler and London Calculation.—The first attack on the solution of the above equation was made by Heitler and London[2] who saw in the known wave functions of the separate atoms a means for approximating the wave function of the molecule.   If the nuclei are far apart, it is not difficult to see that we shall have a good approximate solution of the equation if we write $\psi = f_1 g_2$, where f_1 is an atomic wave

---

[1] D. R. HARTREE, *Proc. Cambridge Phil. Soc.* **24**, 89 (1928).   In this system the fundamental units are as follows:
    Unit of length $= a_0 = h^2/4\pi^2\mu e^2$-radius of innermost Bohr orbit for the hydrogen atom.
    Unit of mass $= \mu =$ mass of an electron.
    Unit of charge $= e =$ charge of an electron.
    Unit of time $= h^3/8\pi^3\mu e^4$.
    Unit of action $= h/2\pi$.
    Unit of energy $= e^2/a_0 = 2 \times$ (ionization energy of normal hydrogen atom).
    The reader is referred to Condon and Shortley, *Theory of Atomic Spectra*, Cambridge and New York, 1935, for a discussion of these units.
[2] W. HEITLER and F. LONDON, *Zeits. f. Physik* **44**, 455 (1927).

function for electron 1 on nucleus A, and g_2 is an atomic wave function for electron 2 on nucleus B. In this case,

$$-\left(\frac{1}{2}\nabla_1{}^2 + \frac{1}{a_1}\right)f_1 = E_1 f_1,$$

$$-\left(\frac{1}{2}\nabla_2{}^2 + \frac{1}{b_2}\right)g_2 = E_2 g_2,$$

where E_1 and E_2 are atomic energy-values. It follows from these relations that

$$[H - (E_1 + E_2)]f_1 g_2 = \left(\frac{1}{r} + \frac{1}{s} - \frac{1}{a_2} - \frac{1}{b_1}\right)f_1 g_2.$$

If r is much larger than the effective radius of the hydrogen atom in either of the states under consideration, the quantity $\left(\dfrac{1}{r} + \dfrac{1}{s} - \dfrac{1}{a_2} - \dfrac{1}{b_1}\right)$ will always be very small when $f_1 g_2$ is relatively large and *vice versa*. Hence the integrals $\displaystyle\int_\infty f_1{}^* g_2{}^* [H - (E_1 + E_2)]f_1 g_2 \, d\tau$ and

$$\int_\infty \Big| [H - (E_1 + E_2)]f_1 g_2 \Big|^2 d\tau$$

will be very small and $f_1 g_2$ must be a good approximate solution of (52·1) for the energy value $E = E_1 + E_2$.

The function $f_1 g_2$ is an eigenfunction of the operator

$$H_0 \equiv -\frac{1}{2}(\nabla_1{}^2 + \nabla_2{}^2) - \frac{1}{a_1} - \frac{1}{b_2},$$

and it would be possible to expand an arbitrary solution of the molecular problem in terms of the complete system of eigenfunctions of this operator. Thus the molecular problem could be dealt with by the conventional perturbation theory of Secs. 47 and 48. It is easy to see, however, that this mode of attack is entirely impracticable. Consider the function $f_2 g_1$ which is identical with $f_1 g_2$ except that it places electron 2 on A in the state previously occupied by 1 and the electron 1 on B in the state previously occupied by 2. $f_2 g_1$ is an eigenfunction of the operator $H_0' \equiv -\dfrac{1}{2}(\nabla_1{}^2 + \nabla_2{}^2) - \dfrac{1}{a_2} - \dfrac{1}{b_1}$, and is as good an approximate solution of (52·1) as $f_1 g_2$. Inspection shows, however, that this new function $f_2 g_1$ is no good at all as an approximate eigenfunction of H_0. This means that expansions of the eigenfunctions of H in terms of those of H_0 will not ordinarily converge with any reasonable degree of rapidity, and that the perturbation theory of Sec. 48 with H_0 as the unperturbed operator will not be useful.

We may reasonably assume, however, that fairly good approximations of the low-energy eigenfunctions of molecular hydrogen can be built up out of the low-energy eigenfunctions of the two symmetrical operators H_0 and H_0'. This is the essential feature of the Heitler and London method. The details are most readily understood as an example of the variational method, though that was not the point of view of the original work. Since there is a wide gap between the energy of the normal state of the hydrogen atom and the energy of the first excited state, it seemed reasonable to use only the normal states of the two atoms as material for the construction of the normal state of the molecule. We accordingly introduce the functions

$$\varphi_1 = f_1 g_2, \qquad \varphi_2 = f_2 g_1,$$

in which f_1, f_2, g_1, g_2 are now identified with normalized normal-state atomic eigenfunctions. In Hartree atomic units these are (*cf.* Sec. 29c)

$$f_1 = \pi^{-\frac{1}{2}} e^{-a_1}, \qquad g_1 = \pi^{-\frac{1}{2}} e^{-b_1},$$
$$f_2 = \pi^{-\frac{1}{2}} e^{-a_2}, \qquad g_2 = \pi^{-\frac{1}{2}} e^{-b_2}.$$

The functions φ_1 and φ_2 are not quite orthogonal and hence we must use Eqs. (51·18) and (51·19) in applying the variational method.

Let S denote the scalar product of φ_1 and φ_2. Since these functions are both real, S is real. Let I_1 and I_2 denote the integrals

$$I_1 = \int_\infty \varphi_1 \left(\frac{1}{s} - \frac{1}{a_2} - \frac{1}{b_1} \right) \varphi_1 d\tau, \qquad I_2 = \int_\infty \varphi_2 \left(\frac{1}{s} - \frac{1}{a_2} - \frac{1}{b_1} \right) \varphi_1 d\tau.$$

Here $d\tau$ is the element of volume in the six-dimensional coordinate space. Let E_0 denote the ionization energy of the normal hydrogen atom with the numerical value $\frac{1}{2}$ in the Hartree units. Then

$$H(1,1) = H(2,2) = 2E_0 + \frac{1}{r} + I_1,$$

$$H(1,2) = H(2,1) = \left(2E_0 + \frac{1}{r} \right) S + I_2,$$

$$\sigma(1,1) = \sigma(2,2) = 1,$$
$$\sigma(1,2) = \sigma(2,1) = S.$$

The secular equation (51·19) becomes

$$\begin{vmatrix} H(1,1) - E & H(1,2) - ES \\ H(1,2) - ES & H(1,1) - E \end{vmatrix} = 0$$

and has the solutions

$$E' = \frac{H(1,1) + H(1,2)}{1 + S} = 2E_0 + \frac{1}{r} + \frac{I_1 + I_2}{1 + S}, \qquad (52\cdot2)$$

$$E'' = \frac{H(1,1) - H(1,2)}{1 - S} = 2E_0 + \frac{1}{r} + \frac{I_1 - I_2}{1 - S}. \qquad (52\cdot3)$$

Let the corresponding relative extremals of Q be

$$\psi' = \xi_1'\varphi_1 + \xi_2'\varphi_2,$$

and

$$\psi'' = \xi_1''\varphi_1 + \xi_2''\varphi_2,$$

respectively. From Eqs. (51·18) we readily derive

$$\xi_1' = \xi_2', \qquad \xi_1'' = -\xi_2''.$$

The normalized solutions are accordingly

$$\psi' = [2(1 + S)]^{-\frac{1}{2}}(\varphi_1 + \varphi_2), \qquad \psi'' = [2(1 - S)]^{-\frac{1}{2}}(\varphi_1 - \varphi_2). \quad (52·4)$$

I_1 is called a "Coulomb" integral as it gives the mean value of the Coulomb energy $\dfrac{1}{s} - \dfrac{1}{a_2} - \dfrac{1}{b_1}$ for the state φ_1. Because the coordinates of the two electrons are exchanged when we pass from one of the two
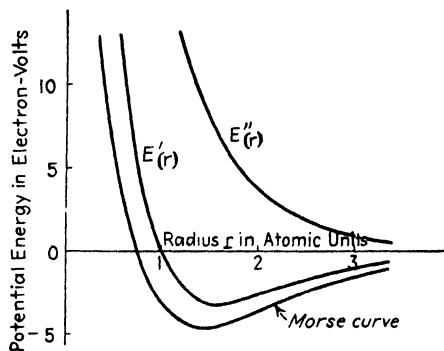
Fig. 22.—Theoretical potential-energy curves for the lowest states of the H_2 molecule compared with a Morse-type curve fitted to the empirical data.

product wave functions which appear in the integrand to the other, I_2 is called an "exchange" integral. The relation of the functions ψ', ψ'' to the Pauli principle, and the interpretation of the energy difference $E' - E''$, approximately determined by I_2, as a by-product of a resonance phenomenon involving the exchange of electrons between the nuclei, is discussed in Sec. 65b.

The integrals S, I_1, and I_2, are, of course, six-dimensional and their evaluation is somewhat troublesome. It has been carried through, however, in part by Heitler and London, and in part by Sugiura[1] to whose paper the reader is referred for the explicit formulas. Figure 22 shows the curves for E' and E'' plotted against the nuclear separation r and for comparison an approximate curve for the apparent potential energy of the hydrogen molecule in its normal state computed backward from the experimental band-spectrum data.[2] From the graph it will

[1] Y. Sugiura, *Zeits. f. Physik* **45**, 484 (1927).
[2] *Cf.* R. S. Mulliken, *Rev. Mod. Phys.* **4**, 73–78 (1932).

be seen that the curve for E' is in reasonable harmony with the empirical data considering the roughness of the calculation, while the E'' curve represents a type of interatomic interaction in which the force between the atoms is apparently repulsive at all distances.[1] The existence of such "repulsive" states for potentially molecule-forming atom pairs was unsuspected before the theoretical work of Heitler and London but is now known to play an essential part in the production of the continuous spectrum of H_2 and in related phenomena.

Instead of deriving the minimizing functions ψ', ψ'' and their energies directly from the non-orthogonal initial approximations φ_1, φ_2 by means of Eqs. (51·18) and (51·19), we might equally well have used φ_1, φ_2 to build up a pair of orthogonal initial approximations which were eigenfunctions of the symmetry integrals of the problem.

Let us identify the internuclear axis with the z axis of a Cartesian-coordinate system and locate the xy plane halfway between the two nuclei. The system has an axial symmetry similar to that met with in Secs. 49 and 50. The group of the Schrödinger equation includes rotations about the z axis, the three reflection operators R_x, R_y, R_z, and the permutation operator P which performs the transformation

$$x_1 \rightarrow \tilde{x}_1 = x_2, \qquad x_2 \rightarrow \tilde{x}_2 = x_1,$$
$$y_1 \rightarrow \tilde{y}_1 = y_2, \qquad y_2 \rightarrow \tilde{y}_2 = y_1,$$
$$z_1 \rightarrow \tilde{z}_1 = z_2, \qquad z_2 \rightarrow \tilde{z}_2 = z_1.$$

All integrals are dependent on $\mathcal{L}_z$, R_x, R_y, R_z and P. But the only non-commuting pairs in this set are $\mathcal{L}_z$, R_x and $\mathcal{L}_z$, R_y. It follows that H combines with each of the three dynamical variables $\mathcal{L}_z$, R_z, R_y to form a complete set of commuting independent dynamical variables. Since R_x has but two eigenvalues, the maximum statistical weight of an energy level is 2.

The dynamical variables $\mathcal{L}_z{}^2$, P, R_z, K are all functions of H since each commutes with every other integral of H. It follows that every discrete energy level of Eq. (52·1) is correlated with a definite eigenvalue of each of the above observables. The eigenvalues of $\mathcal{L}_z{}^2$, P, K are used in the classification of the electronic states of molecular hydrogen. Thus states having the eigenvalues 0, $h^2/4\pi^2$, $2^2h^2/4\pi^2$, $\cdots$ for $\mathcal{L}_z{}^2$ are designated as Σ, Π, Λ, $\cdots$ states respectively. States having the eigenvalues $+1$, -1 for K are designated as *even* and *odd* states. States having the eigenvalues $+1$, -1 for P (*cf.* Sec. 40d) are said to be symmetrical and antisymmetrical, respectively, with respect to an interchange of electrons. When the spin is taken into account each of the

---

[1] Actually, there is a weak attraction at large distances which may be identified with the Van der Waals force of classical gas theory. This is derivable by higher order approximations. The corresponding minimum of E'' is too shallow to permit the formation of molecules of normal stability.

SEC. 52]

antisymmetrical states splits into three substates, whereas no such splitting occurs in the case of the symmetrical states. Consequently the symmetrical states are called "singlets" and the antisymmetrical ones "triplets." The symbol $^3\Pi_g$ denotes an even (*gerade*) triplet state for which $\mathfrak{L}_z{}^2 = \dfrac{h^2}{4\pi^2}$, while $^1\Sigma_u$ denotes an odd (*ungerade*) singlet state for which $\mathfrak{L}_z{}^2 = 0$.

In seeking solutions of Eq. (52·1) we can always deal with eigenfunctions of a specific set of eigenvalues of $\mathfrak{L}_z{}^2$, P, and K. The functions φ_1 and φ_2 are both eigenfunctions of $\mathfrak{L}_z{}^2$ with the eigenvalue 0. Hence any linear combination will be the same. On the other hand, for these very simple forms the operators P and K are equivalent and yield

$$K\varphi_1 = \varphi_2, \qquad K\varphi_2 = \varphi_1.$$

It follows that

$$K(\varphi_1 + \varphi_2) = \varphi_1 + \varphi_2,$$
$$K(\varphi_1 - \varphi_2) = -(\varphi_1 - \varphi_2).$$

Thus the sum and difference of φ_1 and φ_2 are simultaneous eigenfunctions of $\mathfrak{L}_z{}^2$, K, P as desired. Incidentally they are orthogonal. Normalization yields ψ' and ψ'' as the only suitably symmetrized approximate solutions of (52·1) which can be built up out of the atomic wave functions f and g. The application of the variational method to these two functions then reduces to the evaluation of the corresponding energies:

$$E' = \int \psi' H \psi' d\tau = \frac{1}{2(1+S)} \int (\varphi_1 + \varphi_2) H(\varphi_1 + \varphi_2) d\tau$$

$$= \frac{1}{2(1+S)}[H(1,1) + H(2,2) + H(1,2) + H(2,1)] = \frac{H(1,1) + H(1,2)}{1+S},$$

$$E'' = \int \psi'' H \psi'' d\tau = \frac{H(1,1) - H(1,2)}{1-S}.$$

52c. The Method of James and Coolidge.—A decided improvement on the calculation of Heitler and London was effected by Wang,[1] who introduced a single nonlinear parameter λ in the wave function and was able to reduce the previous discrepancy of 1.58 electron-volts between the experimental and theoretical values of the dissociation energy to 0.96 electron-volt.

Of the numerous attempts to make still more precise computations the most successful is that of James and Coolidge,[2] who have used a method of attack resembling that previously used by Hylleraas[3] on the helium atom. These authors, abandoning all attempt to employ atomic

[1] S. WANG, *Phys. Rev.* **31**, 579 (1928).
[2] H. M. JAMES and A. S. COOLIDGE, *J. Chem. Phys.* **1**, 825 (1933), **3**, 129 (1935).
[3] E. A. HYLLERAAS, *Zeits. f. Physik* **54**, 347 (1929).

wave functions as a basis for the treatment of the molecular problem, have introduced a system of non-orthogonal base functions involving products of powers of the elliptic coordinates of the two electrons and of the distance between the electrons. Their approximating wave functions were then built up as linear combinations of the base functions, and Eqs. (51·18) and (51·19) were employed to determine the best values of the coefficients together with the corresponding energy. Using a 13-term wave function and adding a small extrapolated correction for additional terms not actually worked out, they obtained a computed dissociation energy of 4.454 ± 0.013 electron-volts as compared with the recent experimental value 4.455 ± 0.008 electron-volts.

CHAPTER XII

QUANTUM STATISTICAL MECHANICS AND THE EINSTEIN TRANSITION PROBABILITIES

53. QUANTUM STATISTICAL MECHANICS

53a. The General Theory of Perturbations Which May Involve the Time.—Up to this point we have been concerned almost exclusively with problems in which the Hamiltonian operator H is assumed to be independent of time and in which the energy is conserved. In such cases we obtain a unique physically admissible solution of the second Schrödinger equation

$$H\Psi = -\frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} \qquad (53\cdot1)$$

from every physically admissible initial function $\Psi_0(x) = \psi(x)$ by analyzing ψ into eigenfunctions of H and setting $\Psi(x,t)$ equal to the expansion [cf. Eq. (35·1)]

$$\Psi(x,t) = \sum_n \xi_n \psi_n e^{-\frac{2\pi i}{h}E_n t} + \sum_j \int \xi_j(E)\psi_j(E,x) e^{-\frac{2\pi i E t}{h}} dE. \qquad (53\cdot2)$$

Our perturbation theory has been concerned with the effect of a modification in H on its eigenvalues and eigenfunctions.

In the present chapter we shall be concerned with problems in which a different type of perturbation theory is useful—a type which can be applied even when H depends explicitly on the time. Problems of this type have been treated by Schrödinger,[1] Dirac,[2] Born,[3] and Slater.[4] In such cases the system is nonconservative. The expansion (53·2) with constant ξ's no longer gives a solution of (53·1) if the ψ_n's and E_n's are identified with the instantaneous eigenfunctions and eigenvalues of H. It then becomes convenient to study the variation of $\Psi(x,t)$ in time with the aid of an expansion in terms of eigenfunctions of a *constant* approximate Hamiltonian H_0. Such an expansion is possible at every instant of time if the eigenfunctions of H_0 form a complete set. If we identify the functions ψ_n and the energies E_n of (53·2) with the eigen-

---

[1] E. Schrödinger, *Ann. d. Physik* (4) **81**, 109 (1926); **83**, 956 (1927).

[2] P. A. M. Dirac, *Proc. Roy. Soc.* **A112**, 661 (1926).

[3] M. Born, *Zeits. f. Physik* **40**, 167 (1926). *Cf.* also Born and Jordan, *Elementare Quantenmechanik*, Sec. 45, Berlin, 1930.

[4] J. C. Slater, *Proc. Nat. Acad. Sci.* **13**, 7, 104 (1927).

functions and eigenvalues, respectively, of H_0, the function $\Psi(x,t)$ defined by the equation will become a solution of (53·1) provided that the coefficients ξ_n and $\xi_j(E)$, hitherto treated as constants, are required to vary in a suitable manner with the time. The purpose of our present theory is to study the dependence of the ξ's on the time by successive approximations. The procedure was suggested independently by Dirac and others and is often called the Dirac method of the *variation of constants*. It is applicable even when the exact Hamiltonian H is time-free.

From (35·2) or (36·77) it follows that the expansion coefficients ξ_n, $\xi_j(E)$ define the expectation values of the corresponding eigenvalues of the observable H_0 for an assemblage of systems in a pure state with the wave function $\Psi(x,t)$. These expectation values vary with time *as if* transitions from one eigenvalue of H_0 to another were taking place. Application of this method to the problem of the interaction between matter and radiation leads to a quantum-mechanical theory of the Einstein transition probabilities for radiative processes.

In working out the details of the theory we shall use a notation appropriate to the case in which H_0 has a purely discrete spectrum. As previously noted we can always replace the continuous spectrum by a dense discrete spectrum if we make a suitable modification of the boundary conditions or of the natural operator H_0; or the theory here developed can be extended by the introduction of the generalized matrices of Sec. 44d.

As a first step toward the solution of the perturbation problem we reduce Eq. (53·1) to a convenient matrix form. Replacing H by

$$H_0 + \lambda H_1$$

and substituting the expansion $\sum_n \psi_n \xi_n e^{-\frac{2\pi i}{h} E_n t}$ for Ψ, we obtain

$$\sum_n \sum_m \psi_m \lambda H_1(m,n) \xi_n e^{-\frac{2\pi i}{h} E_n t} = -\frac{h}{2\pi i} \sum_n \psi_n \dot{\xi}_n e^{-\frac{2\pi i}{h} E_n t}.$$

The formation of the scalar product of each side of this equation with $\psi_k(x)$ yields

$$\lambda \sum_n H_1(k,n) e^{\frac{2\pi i}{h}(E_k - E_n)t} \xi_n = -\frac{h}{2\pi i} \dot{\xi}_k; \qquad k = 1, 2, 3, \cdots \quad (53\cdot3)$$

Let $\overrightarrow{\xi(t)}$ denote the vector matrix whose components are $\xi_1(t)$, $\xi_2(t)$, $\cdots$ and let $\mathbf{H}_1^{(t)}$ denote the matrix whose typical element is

$$H_1^{(t)}(m,n) = H_1(m,n) e^{\frac{2\pi i}{h}(E_m - E_n)t} = \left(H_1 \psi_n e^{-\frac{2\pi i}{h} E_n t}, \psi_m e^{-\frac{2\pi i}{h} E_m t} \right).$$

Then the system of equations (53·3) can be written in the form[1]

$$\lambda \mathbf{H}_1^{(t)} \overrightarrow{\xi(t)} = -\frac{h}{2\pi i} \frac{d\overrightarrow{\xi(t)}}{dt}. \tag{53·4}$$

Integration with respect to t yields

$$\overrightarrow{\xi(t)} - \overrightarrow{\xi(0)} = -\frac{2\pi i \lambda}{h} \int_0^t \mathbf{H}_1^{(t')} \overrightarrow{\xi(t')} dt', \tag{53·5}$$

where $\overrightarrow{\xi(0)}$ denotes the constant vector matrix made up of the constant values of the ξ_k's for negative values of t. Let us now expand $\overrightarrow{\xi(t)}$ in a power series in λ:

$$\overrightarrow{\xi(t)} = \overrightarrow{\xi(0)} + \sum_{p=1}^{\infty} \lambda^p \overrightarrow{\xi(t)}^{(p)}, \tag{53·6}$$

where $\overrightarrow{\xi(t)}^{(p)}$ is a vector matrix to be determined. Substituting this expansion into Eq. (53·5) and equating the coefficients of like powers of λ, we obtain

$$\overrightarrow{\xi(t)}^{(p)} = -\frac{2\pi i}{h} \int_0^t \mathbf{H}_1^{(t')} \overrightarrow{\xi(t')}^{(p-1)} dt', \tag{53·7}$$

or

$$\xi(t)_k^{(p)} = -\frac{2\pi i}{h} \int_0^t \sum_n H_1(k,n) e^{\frac{2\pi i}{h}(E_k - E_n)t'} \xi(t')_n^{(p-1)} dt'. \tag{53·8}$$

These equations complete the determination of $\overrightarrow{\xi(t)}$ and hence of $\Psi(t)$, thus solving our perturbation problem. It is advantageous, however, to restate the theory in a slightly altered form. To this end we recall that the unitary operator

$$T(t) = e^{-\frac{2\pi i}{h} tH}$$

transforms $\Psi(q,0)$ into $\Psi(q,t)$. It has a unitary matrix $\mathbf{T}(t)$ which transforms the vector $\overrightarrow{\xi(0)}$ into the vector $\overrightarrow{\mathbf{X}(t)}$ of footnote 1.

---

[1] Let $\overrightarrow{\mathbf{X}(t)}$ denote the vector obtained from $\overrightarrow{\xi(t)}$ by multiplying each component by the corresponding phase factor $e^{-\frac{2\pi i}{h} E_n t}$. Then it is a simple matter to derive the equation

$$\overrightarrow{\mathbf{H}\mathbf{X}} = -\frac{h}{2\pi i} \frac{d\overrightarrow{\mathbf{X}}}{dt},$$

which forms an exact matrix parallel to Eq. (53·1).

$T(t)$ can be factored into the product of the unitary commuting operators $e^{-\frac{2\pi i t}{h}H_0}$ and $e^{-\frac{2\pi i \lambda t}{h}H_1}$. Hence $T(t)$ can be factored into the product of the unitary matrices $T_0(t)$ and $T_1(t)$ defined by

$$T_0(t)[m,n] = \int \psi_m{}^* e^{-\frac{2\pi i t}{h}H_0}\psi_n d\tau = e^{-\frac{2\pi i E_n t}{h}}\delta_{mn},$$

$$T_1(t)[m,n] = \int \psi_m{}^* e^{-\frac{2\pi i \lambda t}{h}H_1}\psi_n d\tau.$$

We readily verify that

$$\overrightarrow{X(t)} = T_0(t)\overrightarrow{\xi(t)} = T_0(t)T_1(t)\overrightarrow{\xi(0)}.$$

Since $T_0(t)$ is unitary, we can multiply by its reciprocal as an antecedent factor to obtain

$$\overrightarrow{\xi(t)} = T_1(t)\overrightarrow{\xi(0)}. \tag{53.9}$$

Thus the determination of $\overrightarrow{\xi(t)}$ is reduced to that of evaluating $T_1(t)$.

Instead of attempting to evaluate $T_1(t)$ directly from the corresponding operator we introduce this expression for $\overrightarrow{\xi(t)}$ into Eq. (53·5) and obtain

$$T_1(t)\overrightarrow{\xi(0)} = \overrightarrow{\xi(0)} - \frac{2\pi i \lambda}{h}\left(\int_0^t H_1^{(t')}T_1(t')dt'\right)\overrightarrow{\xi(0)}. \tag{53.10}$$

Since this matrix equation should hold for every normalized vector $\overrightarrow{\xi(0)}$, we conclude that

$$T_1(t) = I - \frac{2\pi i \lambda}{h}\int_0^t H_1^{(t')}T_1(t')dt', \tag{53.11}$$

where I is the unit matrix. To solve for $T_1(t)$ we again develop in power series in λ, remembering that T_1 must reduce to I when $\lambda = 0$. Then

$$T_1(t) = I + \sum_{p=1}^{\infty}\lambda^p T_1(t)^{(p)}, \tag{53.12}$$

if we define $T_1(t)^{(p)}$ by the equation

$$T_1(t)^{(p)} = -\frac{2\pi i}{h}\int_0^t H_1^{(t')}T_1(t')^{(p-1)}dt', \tag{53.13}$$

where $T_1(t)^{(0)} = I$.

The advantage of this form of the theory is that it gives a convenient expression for the so-called "transition probability" from one eigenstate of the observable H_0 to another. Let us consider a case in which the

perturbing Hamiltonian H_1 vanishes outside the time interval $0 < t < \theta$. The transition probability $\Phi_{n \to m}$ from the state ψ_n to the state ψ_m is then defined by Born as the probability that the system will be found in the state ψ_m after the application of the perturbation when the initial state is definitely known to be ψ_n. In other words $\Phi_{n \to m}$ is the value of $|\xi_m(\theta)|^2$ when $|\xi_n(0)|$ is unity and the other components of $\overrightarrow{\xi(0)}$ are zero (*cf.* Sec. 36e). Let us simplify the notation by introducing the symbol $\mathbf{F}$ for $\mathbf{T}_1(\theta)$. Then

$$
\begin{aligned}
\Phi_{n \to m} &= |F(m,n)|^2 = |T_1(\theta)[m,n]|^2 \\
&= \delta_{mn}\{1 + \lambda[F^{(1)}(m,n) + F^{(1)}(m,n)^*] + \lambda^2[F^{(2)}(m,n) \\
&\quad + F^{(2)}(m,n)^*] + \cdots \} + \lambda^2|F^{(1)}(m,n)|^2 + \cdots . \quad (53 \cdot 14)
\end{aligned}
$$

This expression can be reduced still further by means of the relation

$$
\sum_k F(m,k)F(n,k)^* - \delta_{mn} = 0, \qquad (53 \cdot 15)
$$

which expresses the unitary character of $\mathbf{F}$. Expanding in powers of λ and equating the coefficient of each power to zero, we obtain

$$
\left.
\begin{aligned}
F^{(1)}(m,n) + F^{(1)}(n,m)^* &= 0, \\
F^{(2)}(m,n) + F^{(2)}(n,m)^* + \sum_k F^{(1)}(m,k)F^{(1)}(n,k)^* &= 0.
\end{aligned}
\right\} \quad (53 \cdot 16)
$$

In Eq. $(53 \cdot 14)$ it is evidently legitimate to replace n by m in the coefficient of δ_{mn}. Then, using $(53 \cdot 16)$, we find

$$
\Phi_{n \to m} = \delta_{mn}\left[1 - \lambda^2 \sum_k |F^{(1)}(m,k)|^2 + \cdots \right] + \lambda^2|F^{(1)}(m,n)|^2 + \cdots .
$$
$$
(53 \cdot 17)
$$

Thus the transition probabilities are determined to terms of the second degree in λ from the first-degree term of the matrix $\mathbf{F} = \mathbf{T}_1(\theta)$.

The transition probability $\Phi_{m \to n}$ is to be distinguished from the Einstein transition probabilities which govern the statistical averages of the quantum "jumps" of a chaotic assemblage of atomic or molecular systems in a natural radiation field. The Einstein transition probabilities will be discussed in Sec. 54. Here we note that it is not possible without proof to compute the probability of a final state ψ_n for an arbitrary initial state simply by summing the products of the various $|\xi_m(0)|^2$ values by the corresponding Born transition probabilities.

53b. The Adiabatic Theorem.[1]—Another important type of perturbation problem involving the time is that in which the Hamiltonian $H(t)$ does not return to its original form after the time θ and in which the total

[1] *Cf.*, *e.g.*, M. BORN and V. FOCK, *Zeits. f. Physik* **51**, 165 (1928).

change in the time θ is appreciable. In this case the analysis of the perturbed wave functions into eigenfunctions of the unperturbed Hamiltonian is pointless and we have to consider instead the analysis of $\Psi(x,t)$ into eigenfunctions of the instantaneous Hamiltonian as given by the equation

$$H(t)\varphi_n(x,t) = E_n(t)\varphi_n(x,t). \qquad (53\cdot18)$$

In order to obtain a set of orthogonal eigenfunctions of this equation which are uniquely defined except for constant phase-factors it is necessary to introduce the auxiliary condition $(\partial\varphi_n/\partial t, \varphi_n) = 0$ which may be compared with (47·10).

A case of special importance is that in which the operator $H(t)$ varies adiabatically, *i.e.*, very slowly with the time. The slow applications of external electric and magnetic fields to atomic and molecular systems are examples in point. Here the quantum-mechanical extension of Ehrenfest's adiabatic theorem applies. This states that if the Hamiltonian changes from the initial form H_0 to the final form H_1 in the time θ, and if the value of θ is allowed to approach infinity while the rate at which the Hamiltonian changes approaches zero, the transition probability from the state m to the state n approaches zero. In other words, solutions of (53·18) subject to the above-mentioned auxiliary condition become solutions of $H(t)\Psi = -\dfrac{h}{2\pi i}\dfrac{\partial\Psi}{\partial t}$ as the rate of change of $H(t)$ approaches zero. In case we have to do with the application of an external field there is a gradual transformation of the eigenvalues and eigenfunctions of H, but no forced jumps from one state as defined by (53·18) to another. The theorem is proved—with suitable restrictions on $H(t)$—by Born and Fock.[1]

53c. The Fundamental Problems of Quantum Statistical Mechanics.

As a first application of the general theory of Sec. 53a we shall consider the fundamental features of the quantum-mechanical form of statistical mechanics. This branch of physical science has to do with the statistical properties of assemblages of identical systems not so prepared as to have a definite wave function but having a definite temperature. If we approach the problem from the standpoint of Pauli[2] our treatment takes a form which parallels the usual developments of classical statistical mechanics and at the same time affords an important illustration of our perturbation theory.

---

[1] *Loc. cit.*

[2] W. PAULI, JR., in the Sommerfeld Festschrift, *Probleme der Modernen Physik*, p. 31, Leipzig, 1928; *Handbuch d. Physik*, Geiger u. Scheel, 2d ed., Band XXIV/1, Kap. 2, Ziff. 10, 1933. *Cf.* also P. Jordan *Statistische Mechanik auf Quantentheoretische Grundlage*, Braunschweig, 1933.

Experiment shows that the statistical properties[1] of a large assemblage of independent identical microscopic, or macroscopic, systems (*i.e.*, a Gibbsian assemblage) which have been "aged" in a thermostat at a definite temperature T for a sufficient length of time usually become constant and independent of the initial state of the assemblage. The ultimate state is then defined to be one of thermodynamic equilibrium at the temperature T. By erasing all vestiges of the initial state the thermostat acts as a history-destroying device. To be sure there are numerous cases in which this function is imperfectly performed. In such cases the state of true thermodynamic equilibrium, or maximum entropy, is not reached in any measurable time at moderate temperatures. We may restrict the discussion for the present, however, to systems for which thermodynamic equilibrium is actually attainable.

If the systems under consideration are of large scale, the dispersion of the results for certain types of measurement, notably that of energy, becomes negligible compared with experimental error when the aging process has been carried to completion. Such quasi-unique properties of the individual system in an assemblage in thermal equilibrium are often called *normal* properties of the assemblage. In any case, however, there exist other properties which exhibit measurable fluctuations, but with distribution functions which have definite forms for states of thermal equilibrium.

From the standpoint of quantum mechanics the properties of the most general Gibbsian assemblage are postulated to be those of an appropriate mixture of pure states. The experimental facts then inform us that an arbitrary mixture when subjected to interaction with a thermostat for a sufficient length of time approaches a standard state with properties uniquely defined by the temperature, the volume of the container, or other externally variable parameters. An assemblage in the ultimate state of thermal equilibrium will be referred to hereafter as a *thermostat assemblage*. The fundamental problems of quantum statistical mechanics are (**a**) *the explanation of the approach to thermal equilibrium* and (**b**) *the development of a detailed description of the corresponding thermostat assemblage.* Solution of these problems paves the way for a quantum-mechanical derivation of the laws of thermodynamics and for a quantum-mechanical theory of the properties of matter in bulk. In fact a solution of the second problem is a prerequisite to any satisfactory application of quantum-mechanical ideas to matter in bulk. In order to test the application of the laws of quantum mechanics to any particular type of system we must have a method for preparing an assemblage of systems of the given type whose statistical state is known in sufficient detail to permit predictions regarding measurements we are able to make. In

---

[1] By "statistical properties" we mean the totality of the distribution functions representing the results of all types of statistical measurement.

the case of microscopic systems we can prepare assemblages which bear some resemblance to pure states in order to test the theory. On account of the appalling complexity of macroscopic systems the preparation of approximate pure states is for them out of the question. We are accordingly driven to the use of thermostat assemblages as the "known" starting point for tests of the theory.

It is unfortunately impossible to give definite form to the equations of motion of a system, or an assemblage, subject to unknown perturbations from a thermostatic container. Hence it is customary to approach problem (a) by way of a study of the idealized case of an assemblage of isolated systems. Fortunately experiment indicates that the tendency toward thermodynamic equilibrium exists for approximately isolated systems as well as for systems in intimate contact with thermostats. In fact it is a familiar extrapolation of experimental results to say that an assemblage of isolated large-scale systems started in any state with energies restricted to the macroscopically small range between E and $E + \Delta E$ will in time reach an apparently constant state experimentally indistinguishable from one of thermal equilibrium with a thermostat of temperature T appropriate to E. Our first theoretical step will then be to verify from the principles of quantum mechanics that the state of such an assemblage after a sufficient length of time will become experimentally indistinguishable from a constant state the same for all initial conditions consistent with the given energy E. This constant asymptotic state is used as a first approximation to the desired thermostat assemblage. To get a better approximation we can imagine each of the systems A under consideration united with a corresponding thermostat B into a single isolated system C. Using an approximate thermostat assemblage for C we can work out the statistical state of the A systems in contact with thermostats. If the heat capacity of the A systems is small compared with that of the thermostats, the resulting statistical state is independent of the exact energy distribution function adopted for the assemblage of supersystems C. Thus the statistical state of the A systems obtained by this second approximation in the asymptotic case where the ratio of the heat capacities is $1/\infty$ can be adopted as a satisfactory model for the thermostat assemblage. It is in fact the *canonical* assemblage of Gibbs.

53d. The Conventional Characterization of a Chaotic Assemblage.— It will be convenient to restrict the discussion for the present to systems which can be resolved into a sum of parts (*e.g.*, molecules) whose energies are large on the average compared with their mutual energies. Let H denote the Hamiltonian of the system as a whole. It will be resolvable into the sum of terms H_0 and H_1, the former representing the sum of the energies of the individual molecules, and the latter representing the mutual energies which come into play classically in collisions. In treating H_1 as small compared with H we restrict the discussion to

approximations of the ideal-gas state. In order to take into account the finite volume of the container an artificial term must be introduced into the potential-energy function of each molecule which abruptly becomes infinite when that molecule strikes the surface of the vessel. This artificial term takes the place of the much more complicated wall reaction for the case where the gas and the containing vessel are thought of as parts of a single quantum-mechanical system. Owing to the finite volume of the container the spectra of H_0 and H_1 will both be purely discrete, though very dense. If we postulate "rough" walls, the rotation-reflection group will be eliminated from the group of the Schrödinger equation even when the containing vessel is macroscopically spherical. The permutation group yields no degeneracy on account of the Pauli principle. Hence the eigenvalues of H are physically non-degenerate, and we conclude that if the energy of one of the isolated systems were exactly known, it would be in a pure state. On the other hand the eigenvalues of H_0 may be highly degenerate.

In order to study the statistical properties of mixed assemblages it is convenient to make use of the *statistical matrix* first introduced by von Neumann,[1] to characterize such an assemblage. Let $\psi_1, \psi_2, \cdots \psi_n, \cdots$ denote a set of functions which is complete and orthonormal in the configuration space of the system. Let w_s and Ψ_s denote the weight and wave function of the sth pure-state subassemblage of the mixture under consideration. Let $c_{ns}^{(t)}$ denote the Fourier coefficient defined by

$$\Psi_s = \sum_n \psi_n c_{ns}^{(t)}; \qquad c_{ns}^{(t)} = (\Psi_s, \psi_n). \tag{53.19}$$

Let α be any observable. Its expectation value (*cf.* Sec. 41*b*) is

$$\bar{\alpha} = \sum_s w_s (\alpha \Psi_s, \Psi_s). \tag{53.20}$$

Substituting the expansion (53.19), we obtain

$$\bar{\alpha} = \sum_n \sum_m \alpha(m,n) \sum_s w_s c_{ns}^{(t)} c_{ms}^{(t)*},$$

where $\alpha(m,n)$ is the usual matrix element $(\alpha \psi_n, \psi_m)$. Let the Hermitian matrix ϱ be defined by

$$\rho(n,m) = \sum_s w_s c_{ns}^{(t)} c_{ms}^{(t)*}. \tag{53.21}$$

Then

$$\bar{\alpha} = \sum_m \sum_n \alpha(m,n)\rho(n,m) = \text{Spur } \alpha\varrho, \tag{53.22}$$

[1] J. VON NEUMANN, *Gött. Nachr.*, p. 245 (1927); *M.G.Q.*, p. 167. *Cf.* also P. A. M. DIRAC, *Proc. Camb. Phil. Soc.* **25**, 62 (1929).

an invariant of a canonical transformation. The probabilities of different eigenvalues of α can be calculated with equal simplicity from ρ (they are expectation values of suitably defined linear operators) so that ρ characterizes the entire statistical behavior of the assemblage and is properly called its statistical matrix.

In the classical statistical mechanics of Gibbs an assemblage of independent systems is said to be in statistical equilibrium if the expectation value of every function of the coordinates and momenta is constant in time. We accordingly define *statistical equilibrium* for a mixed assemblage in quantum mechanics by the requirement that ρ shall be constant in time. If we use exact eigenfunctions of the complete Hamiltonian H for the ψ_n's, the expansion coefficients c_{ns} take the form

$$c_{ns}^{(t)} = \gamma_{ns} \exp\left[i\left(-\frac{2\pi}{h}E_n t + \theta_{ns}\right)\right],$$

where γ_{ns} and θ_{ns} are real constants. Introducing the abbreviation ν_{mn} for $(E_m - E_n)/h$, we have

$$\rho(n,m) = e^{2\pi i \nu_{mn} t}\sum_s w_s \gamma_{ns}\gamma_{ms}e^{i(\theta_{ms}-\theta_{ns})}.$$

Since the energy levels E_n are nondegenerate, all the off-diagonal elements of ρ are periodic functions of the time and it is necessary for statistical equilibrium that ρ be a diagonal matrix. Reduction of ρ to this form implies either that every component pure-state assemblage shall have a unique energy, or that the values of the phase differences $\theta_{ms} - \theta_{ns}$ shall be uniformly distributed over the range from $-\pi$ to $+\pi$ so that

$\sum_s w_s \gamma_{ns}\gamma_{ms}e^{i(\theta_{ms}-\theta_{ns})}$ vanishes unless $m = n$. Thermostat assemblages

are special cases of statistical equilibrium which we interpret most plausibly by the hypothesis of uniform phase distribution.

Since the average values of the off-diagonal elements of ρ are zero, it is evident that the time-average of the expectation value of α for a sufficiently long time is obtained from (53·22) by neglecting the off-diagonal elements of ρ and α. The initial conditions may be such that these off-diagonal elements make an important contribution to $\bar{\alpha}$ at the time $t = 0$. However, we can be sure that there will be no systematic commensurabilities among the periods of the off-diagonal elements. Hence the distribution of their phases will become chaotic in time. Furthermore, it is safe to assume that in any initial statistical state which it is possible to prepare there will be a large number of off-diagonal elements whose absolute values are of the same order of magnitude as their mean.[1]

---

[1] The Gibbs classical theory indicates that the order of magnitude of the uncertainty in the energy of a system in thermal equilibrium is equal to the average kinetic energy per degree of freedom multiplied by the square root of the total number of degrees of freedom. We cannot hope to determine the energy of a macroscopic system

In general the chaotic distribution of phases among these elements implies a cancellation of contributions to $\bar{\alpha}$. From time to time accidental correlations of the phases of ϱ and α will lead to momentary deviations between the exact value of $\bar{\alpha}$ and that computed from the diagonalized average of ϱ. Similarly the distribution function for the eigenvalues of α will fluctuate about an average form given by the diagonal matrix $\bar{\varrho}$. But *experimentally* haphazard deviations in the distribution function for the eigenvalues of α are not to be distinguished from the deviations in the measured values of α allowed by the average distribution function. Hence we can say that the statistical state of an arbitrary assemblage of isolated systems started off in any practical way becomes—after a time sufficient to insure a chaotic distribution of phases—experimentally indistinguishable from the state of statistical equilibrium whose matrix is $\bar{\varrho}$.

Although use of a matrix scheme for ϱ based on an orthonormal system of eigenfunctions of the complete Hamiltonian H leads at once to a simple criterion for statistical equilibrium analogous to that given by Liouville's theorem in classical theory, this matrix scheme is not the one which makes the closest connection with the language in which statistical mechanics is ordinarily discussed. The eigenfunctions of H are so complicated that we can have little idea of their character in detail. Thus the evaluation of the matrix of any observable α which could be measured would be impossible if a scheme of this type were employed. It is therefore convenient to replace the functions ψ_n of Eqs. (53·19) by eigenfunctions $u_n(x)$ of the *approximate* Hamiltonian H_0. In place of (53·19) we have

$$\Psi_s = \sum_n u_n \xi_{ns}; \qquad \xi_{ns} = (\Psi_s, u_n). \tag{53·23}$$

Let ϱ' denote the statistical matrix in the new scheme with the typical element $\varrho'(n,m) = \sum_s w_s \xi_{ns} \xi_{ms}^*$. If $\mathbf{U}$ is the unitary matrix which transforms $\overrightarrow{c}$ into $\overrightarrow{\xi}$ according to the rule $\overrightarrow{\xi} = \mathbf{U}\overrightarrow{c}$, we have

$$\rho'(n,m) = \sum_l \sum_p U(n,l)\rho(l,p)U^{-1}(p,m). \tag{53·24}$$

---

more accurately. Hence the uncertainty in the total energy of any actual macroscopic system, although small compared with the expectation value, is sure to be astronomically large in comparison with the spacing of the energy levels of the operators H and H_0. Thus the total number of off-diagonal elements of ϱ which can have appreciable values is enormous. We have no experimental means of choosing initial statistical states which are approximate eigenfunctions of H and on the basis of any plausible assumption regarding *a priori* probabilities it is evident that initial states involving anything but a haphazard distribution of absolute values over the array of possible elements of ϱ are extremely improbable. Hence we *postulate* such a haphazard distribution.

The elements of the matrix ϱ' can be divided into two classes consisting respectively of those elements connecting different eigenvalues of H_0 and those connecting different states with the same eigenvalue of H_0. The elements of the former class yield sensibly zero time-averages since these elements are in first approximation harmonic functions of t with frequencies given by the differences between the two correlated eigenvalues of H_0. Elements of the latter class, including the diagonal elements of ϱ', are in first approximation constant, but undergo secular changes in time and do not give zero time-averages. Nevertheless it is customary in statistical mechanics to neglect *all* off-diagonal elements of ϱ'. The discarding of off-diagonal elements which do not average to zero in time is justified by the observation that in the absence of a correlation between the average phases of these elements and the phase of the matrix of any measurable physical quantity α the contribution of these elements to the expectation value of α or any function of α will be small. Assemblages for which it is ordinarily legitimate to neglect the off-diagonal elements of ϱ' are conveniently labeled *chaotic*.

The above approximation is of importance because of the connection it makes with the conception of the stationary state characteristic of the Bohr theory. In that theory it was assumed that every molecule of a gas must spend practically all its time in one or another of the discrete quantized single-energy states which in our present theory we should correlate with the members of a system of orthonormal eigenfunctions of the Hamiltonian of the individual molecule. Transfers from one state to another were assumed to take place discontinuously as a result of collisions and interaction with the ever-present radiation field. The primary problem of statistical theory was then to determine the average number of molecules in each individual molecule energy level for an average sample of gas at a temperature T

The picture behind such a computation is inadmissible from our present point of view. The allowed "microscopic states" of the complete sample of gas in the Bohr theory are defined by particular distributions of the molecules among the allowed states for an individual molecule. The analogues of these states of the macroscopic system in quantum mechanics are pure states of a special type whose wave functions are the products of eigenfunctions of the single-molecule Hamiltonian, or pure states formed from such products by a symmetrization process in accordance with the Pauli principle. These states can be correlated with the functions $u_1, \cdots u_n, \cdots$ of (53·23). Hence the analogue of the Bohr concept of a Gibbsian assemblage of independent identical macroscopic samples of gas is a mixture of pure states whose wave functions are all members of the sequence $u_1, \cdots u_n, \cdots$. The statistical matrix ϱ' for such a mixture would be diagonal. It would not preserve

this form for any finite period of time, however, since the assumed wave functions are not solutions of the appropriate Schrödinger equation,

$$H\psi = -\frac{h}{2\pi i} \cdot \frac{\partial \Psi}{\partial t}.$$

Nevertheless, it is convenient to correlate the state of a macroscopic system as defined by actual experimental conditions with a corresponding mixture of the above type suggested by the Bohr picture. This simply means that we neglect the off-diagonal elements of the ϱ' already shown to be small in the great majority of cases. No approximation is involved in discarding these off-diagonal elements when we are computing the expectation value of a dynamical variable α which commutes with H_0 and whose matrix in the u_n scheme is itself diagonal. There is an approximation, however, in the case of a variable, like the external pressure, which does not commute with H_0. Consider, for example, an assemblage of identical gaseous systems in which the density distribution is very non-uniform. In order to calculate the density distribution from the diagonal elements of ϱ' we have to calculate the density distribution corresponding to each of the orthonormal functions u_n, multiply by the corresponding diagonal element of ϱ', and add. Each of the terms will contribute a roughly uniform density and so the density computed in this way must be uniform. It follows at once that at the moment in question the diagonal elements of ϱ' will not tell the truth, i.e., the nondiagonal elements must play an important part. Such states are, however, essentially ephemeral.

Since the off-diagonal elements of ϱ' are to be neglected, it will be convenient in the remainder of this discussion to adopt the language of the older quantum statistics, interpreting each of the diagonal elements of ϱ' as the probability that an arbitrary member of the assemblage is in the state with the wave function u_n. Let N denote the total number of systems in the whole assemblage and let ρ_n' denote the diagonal element $\rho'(n,n)$. Then $N_n = N\rho_n'$ can be interpreted as the *population* of the state u_n.

53e. Transition Probabilities and Statistical Equilibrium for Chaotic Assemblages.—Before carrying the argument farther it will be well to pause in order to formulate the fundamental assumption of most work in the application of statistical mechanics to special problems.[1] It is necessary to mention the existence of certain approximate integrals of the Hamiltonian H which are usually exact integrals of H_0 and which may be used to divide the wave functions $u_1, u_2, \cdots$ into approximately independent classes which are not mixed to any appreciable degree by the action of the perturbation Hamiltonian H_1 in the time required for

---

[1] *Cf.*, *e.g.*, R. H. FOWLER, *Statistical Mechanics*, 2d ed., §1·4, Cambridge and New York, 1936.

the attainment of approximate statistical equilibrium within each class.[1] In such cases the sum of the diagonal elements of ϱ' associated with each independent class—and therefore the total population of the class—is sensibly constant for the duration of many experiments in which quasi-equilibrium is reached, and the relative values of these sums are fixed by the initial conditions and do not correspond to rigorous statistical or thermodynamic equilibrium. The fundamental assumption mentioned above is now that a satisfactory model thermostat assemblage for a macroscopic system in approximate thermodynamic equilibrium is one with a constant diagonal matrix ϱ' whose nonvanishing diagonal elements are associated with a *single* energy level E_n of the approximate Hamiltonian H_0, all populations associated with any one of the independent classes being equal.

This assumption gives correct answers to the problems to which it is customarily applied, but would not be satisfactory for all imaginable purposes since the energy of a macroscopic system is never uniquely defined in the case of a statistical state which is experimentally realizable. We cannot prove that the matrix ϱ' of an arbitrary assemblage of isolated identical macroscopic systems with an initial energy uncertainty ΔE takes the above form for large values of t, but we can prove a proposition which for practical purposes is equivalent to that. To this end we restrict the discussion to the case in which there is a single independent class of wave functions and divide these functions into groups. Each group consists of wave functions u_n associated with energies ranging from E to $E + \Delta E$, but with a single set of eigenvalues of a suitably chosen set of mutually compatible integrals of the unperturbed motion. Each of these groups has a total population easily derivable from ϱ' by summing corresponding diagonal elements. It is easier to study the secular changes in the total populations of these groups than it is to study the changes in the population of the individual substates. For this purpose we shall set up differential equations for the time rate of change of the group populations from which it follows that for large values of t the average population of the substates in one group must become equal to the average population of the substates in any other. But the different substates in any one group can be reckoned physically indistinguishable so that equality in the group-average populations of the substates is for physical purposes

---

[1] For example, in the case of an experimental measurement of the specific heat of H_2 at low temperatures the transformation of orthohydrogen molecules into para-hydrogen molecules and *vice versa* takes place so slowly that we may assume statistical equilibrium with respect to all other types of transition and at the same time treat the relative amounts of orthohydrogen and parahydrogen as constants of the motion fixed once and for all by the initial conditions. Under these circumstances the different values of the ratio of orthohydrogen to parahydrogen divide the wave functions into independent classes which we regard as immiscible.

equivalent to equality in the individual populations of all substates. The method is that of Dirac[1] and Pauli (*loc. cit.*).

The instantaneous time derivative of any diagonal element of ϱ' is given by

$$\frac{d\rho_m'}{dt} = \sum_s w_s \left(\xi_{ms} \frac{d\xi_{ms}^*}{dt} + \frac{d\xi_{ms}}{dt} \xi_{ms}^* \right)$$

and Eq. (53·3). This instantaneous derivative, however, involves high-frequency terms which cause rapid oscillations in $d\rho_m'/dt$ about a mean value which gives the secular change in ρ_m'. We shall be interested in the mean, or secular, derivatives of the diagonal elements and these are most conveniently computed with the aid of the perturbation theory of Sec. 53a.

In place of (53·17) we have

$$|\xi_{ms}(t)|^2 = \sum_n |F(m,n)|^2 |\xi_{ns}(0)|^2 + \sum_n \sum_{n' \neq n} \{ F(m,n)F(m,n')^* \xi_{ns}(0)\xi_{n's}(0)^*$$
$$+ F(m,n)^* F(m,n') \xi_{ns}(0)^* \xi_{n's}(0) \} \quad (53\cdot25)$$

for the case of an arbitrary initial wave function $\Psi_s(0)$. In order to make progress by our chosen method of attack it is necessary to eliminate the double sum and with it the high-frequency contributions to $|\xi_{ms}(t)|^2$. The elements of this double sum are real quantities which can equally well be positive or negative. They will yield a negligibly small contribution if there are many equally important nonvanishing terms in the sum with a haphazard distribution of phases. It is not difficult to see that this condition should be satisfied by the great majority of the possible choices of the coefficients $\xi_{ns}(0)$ describing pure initial states consistent with the condition that a measurement of the total energy must be certain to yield a value in the basic interval ΔE. It follows that if we multiply the above equation by w_s and sum over all subassemblages of a mixture consistent with the same condition, the contribution of the double sum to the result will nearly always be negligible. Hence in the case of any mixed assemblage which has a claim to consideration as a possible model thermostat assemblage,

$$N_m(t) = N\rho_m'(t) = N \sum_s w_s |\xi_{ms}(t)|^2 = \sum_n |F(m,n)|^2 N_n(0). \quad (53\cdot26)$$

From (53·13) we have

$$\mathbf{F}^{(1)} \equiv \mathbf{T}_1(t)^{(1)} = -\frac{2\pi i}{h} \int_0^t \mathbf{H}_1^{(t')} dt'.$$

[1] P. A. M. DIRAC, *Proc. Roy. Soc.* **A114**, 243 (1927).

Thus

$$F^{(1)}(m,n) = -\frac{2\pi i}{h}\int_0^t H_1(m,n)e^{\frac{2\pi i}{h}(E_m-E_n)t'}dt' = \frac{H_1(m,n)}{E_n - E_m}\left[e^{\frac{2\pi i}{h}(E_m-E_n)t} - 1\right]$$

if $E_m \neq E_n$, and

$$F^{(1)}(m,n) = -\frac{2\pi i t}{h}H_1(m,n)$$

if $E_m = E_n$. Neglecting higher order terms in (53·17) and setting λ equal to unity, we can throw (53·26) into the form

$$N_m(t) = N_m(0) + \sum_{n \neq m} |F^{(1)}(m,n)|^2\{N_n(0) - N_m(0)\}$$

$$= N_m(0) + 4\sum_{n}' \frac{|H_1(m,n)|^2}{(E_n - E_m)^2} \sin^2\pi\nu_{nm}t \{N_n(0) - N_m(0)\}. \quad (53\cdot27)$$

Here ν_{nm} is again used to denote $(E_n - E_m)/h$ and $\sum_n'$ denotes a sum over all values of n not equal to m.

In order to derive a differential equation from (53·27) it is necessary to approximate the sum which it contains by means of an integral. To this end we must examine the matrix elements $H_1(m,n)$ in some detail. The perturbing Hamiltonian operator H_1 gives the mutual potential energy of the molecules. It will accordingly be made up of a sum of terms each of which involves the coordinates of two molecules only. The wave functions u_1, u_2, $\cdots$ will be linear combinations of products of individual molecule functions chosen from an orthonormal set φ_1, φ_2, $\cdots$. These linear combinations will be so selected as to accord with the symmetry requirements of the Pauli principle. Every complete function u_n is defined by a set of non-negative integers $a_1^{(n)}$, $a_2^{(n)}$, $\cdots$ which determine the number of molecules assigned to each of the φ's. The $a^{(n)}$'s are, of course, subject to the restriction $\sum_r a_r^{(n)} = N$. The matrix elements of H_1 between two states with wave functions u_n and u_m can be shown to vanish unless the two corresponding sets of integers $a_r^{(n)}$ and $a_r^{(m)}$ are identical except in the case of four values of r, say α, β, γ, δ. The detailed proof will be supplied by the reader without difficulty after he has examined Secs. 63 and 64. Let $V(1,2)$ denote the part of H_1 which gives the mutual energy of molecules 1 and 2. Each nonvanishing matrix element $H_1(m,n)$ turns out to be a multiple of an integral of the form $\int V(1,2)\varphi_\alpha*(1)\varphi_\beta*(2)\varphi_\gamma(1)\varphi_\delta(2)d\tau_1 d\tau_2$ where $d\tau_1$ is an element of volume in the configuration space of molecule 1 and $d\tau_2$ is a similar element for molecule 2 and

$$a_\alpha^{(m)} = a_\alpha^{(n)} + 1; \quad a_\beta^{(m)} = a_\beta^{(n)} + 1; \quad a_\gamma^{(m)} = a_\gamma^{(n)} - 1; \quad a_\delta^{(m)} = a_\delta^{(n)} - 1.$$

Let us now replace the single index n of the wave function u_n, with a pair of indices, say n and ν, with the following properties: n is to fix the translational energy state of one molecule, while ν determines the complete state of all other molecules and simultaneously the internal state of the one. A group of wave functions $u_{n\nu}$ with a single value of ν and a variety of choices of n can be reckoned macroscopically equivalent. With the double index notation the typical matrix element of H_1 becomes

$$H_1(m,\mu;\; n,\nu) \sim \int V(1,2)\varphi_\alpha{}^*(1)\varphi_\beta{}^*(2)\varphi_\gamma(1)\varphi_\delta(2)d\tau_1 d\tau_2.$$

If we hold m, μ, ν fast and vary n, we change only the translational factor of one of the two functions $\varphi_\gamma(1)$, $\varphi_\delta(2)$. In order that $H_1(m,\mu;\,n,\nu)$ shall be large it is necessary that real or imaginary parts of the product $\varphi_\alpha{}^*(1)\varphi_\beta{}^*(2)\varphi_\gamma(1)\varphi_\delta(2)$ shall be predominantly of one sign in that portion of the configuration space of the molecules 1 and 2 where $V(1,2)$ is large. In other words there must be a constructive interference in configuration space which becomes less and less possible as the total energy of the state $\varphi_\gamma(1)\varphi_\delta(2)$ diverges more and more from that of the state $\varphi_\alpha(1)\varphi_\beta(2)$. Hence the average absolute value of the matrix elements $H_1(m,\mu;\,n,\nu)$, other things being equal, must be a function of $E_{m\mu} - E_{n\nu}$ with a maximum when $E_{m\mu} - E_{n\nu}$ is zero.

Let us now introduce the quantity J which we define as the value of

$$\frac{1}{\Delta\epsilon}\sum_n |H_1(m,\mu;\; n,\nu)|^2 \text{ when the sum is extended over all values of } n \text{ cor-}$$

responding to energies $E_{n\nu}$ in the range $\epsilon - \tfrac{1}{2}\Delta\epsilon < E_{n\nu} < \epsilon + \tfrac{1}{2}\Delta\epsilon$. We assume that we can choose $\Delta\epsilon$ large enough so that the sum includes many terms and small enough so that J changes by a small fraction of itself when ϵ is increased by $\Delta\epsilon$. In accordance with the foregoing discussion we postulate that J is a symmetric function of $\epsilon - E_{m\mu}$ which we write as $J_{\mu\nu}(\epsilon - E_{m\mu})$. The range, say $D\epsilon$, of values of ϵ which contribute appreciably to

$$\int J_{\mu\nu}(\epsilon - E_{m\mu})d\epsilon = \sum_n |H_1(m,\mu;\; n,\nu)|^2$$

is difficult to estimate but can hardly be of order of magnitude greater than that of the average translational energy of one molecule.

A similar average value is defined by

$$K_{\nu\mu}(\epsilon - E_{n\nu}) = \frac{1}{\Delta\epsilon}\sum_{m \text{ in } \Delta\epsilon} |H_1(m,\mu;\, n,\nu)|^2,$$

where the sum is to be extended to all values of m corresponding to energies $E_{m\mu}$ in the interval $\epsilon - \tfrac{1}{2}\Delta\epsilon < E_{m\mu} < \epsilon + \tfrac{1}{2}\Delta\epsilon$.

We rewrite Eq. (53·27) in the double-subscript notation and sum over all values of m for which $E < E_{m\mu} < E + \Delta E$. Introducing the abbreviation $N_\mu(t) = \sum\limits_{m \text{ in } \Delta E} N_{m\mu}(t)$, we obtain

$$N_\mu(t) - N_\mu(0) = -4 \sum_{m \text{ in } \Delta E} N_{m\mu}(0) \sum_\nu \left\{ \sum_n \frac{|H_1(m,\mu; n,\nu)|^2}{(E_{n\nu} - E_{m\mu})^2} \sin^2 \frac{\pi t}{h}(E_{n\nu} - E_{m\mu}) \right\}$$

$$+ 4 \sum_\nu \sum_n N_{n\nu}(0) \left\{ \sum_{m \text{ in } \Delta E} \frac{|H_1(m,\mu; n,\nu)|^2}{(E_{n\nu} - E_{m\mu})^2} \sin^2 \frac{\pi t}{h}(E_{n\nu} - E_{m\mu}) \right\}. \quad (53·28)$$

The two quantities in braces can be replaced by integrals. Thus

$$4 \sum_n \frac{|H_1(m,\mu; n,\nu)|^2}{(E_{n\nu} - E_{m\mu})^2} \sin^2 \frac{\pi t}{h}(E_{n\nu} - E_{m\mu})$$

$$\cong 4 \sum_{\Delta\epsilon} \frac{J_{\mu\nu}(\epsilon - E_{m\mu})}{(\epsilon - E_{m\mu})^2} \sin^2 \frac{\pi t}{h}(\epsilon - E_{m\mu}) \Delta\epsilon$$

$$= \frac{4\pi t}{h} \int J_{\mu\nu}\left(\frac{h\xi}{\pi t}\right) \frac{\sin^2 \xi}{\xi^2} d\xi.$$

The familiar function $\sin^2 \xi / \xi^2$ consists of a series of arches between the zeros at $\pm m\pi$. The area under the dominant central arch is over 90 per cent of the area under the complete graph of the function. Hence we can neglect the other arches in first approximation and can treat $J_{\mu\nu}(h\xi/\pi t)$ as a constant equal to $J_{\mu\nu}(0)$, provided that $t \gg h/D\epsilon$ so that the quantity $|h\xi/\pi t|$ is much less than $D\epsilon$ when $|\xi| < \pi$. Under these circumstances the quantity $J_{\mu\nu}(0)$ can be taken out from under the integral sign and the integral evaluated. The sum reduces to $4\pi^2 t J_{\mu\nu}(0)/h$.

The second expression in braces in (53·28) requires slightly different treatment since we have to sum only over values of m in the interval ΔE. We find

$$\sum_{m \text{ in } \Delta E} \frac{|H_1(m,\mu; n,\nu)|^2}{(E_{n\nu} - E_{m\mu})^2} \sin^2 \frac{\pi t}{h}(E_{n\nu} - E_{m\mu}) \cong \frac{\pi t}{h} \int_\gamma K_{\nu\mu}\left(\frac{h\xi}{\pi t}\right) \frac{\sin^2 \xi}{\xi^2} d\xi$$

$$= \frac{\pi t}{h} K_{\nu\mu}(0) \int_\gamma \frac{\sin^2 \xi}{\xi^2} d\xi,$$

where γ is the range of values of the quantity

$$\xi = \frac{\pi t}{h}(\epsilon - E_{n\nu})$$

which places ϵ in the range ΔE. In other words γ denotes the interval

$$\frac{\pi t}{h}(E - E_{n\nu}) < \xi < \frac{\pi t}{h}(E + \Delta E - E_{n\nu}).$$

The ratio of the width of the central arch of $\sin^2\xi/\xi^2$ to the width of γ is $2h/t\Delta E$. We assume that for the values of t and ΔE under consideration, this quantity is much less than unity. Then the function $\int_\gamma \frac{\sin^2\xi}{\xi^2}d\xi$ will approximate π for nearly all values $E_{n\nu}$ well inside ΔE and will approximate 0 for nearly all values of $E_{n\nu}$ outside γ. The sum we are trying to evaluate is to be interpreted as a multiple of the transition probability from the state $u_{n\nu}$ to the group of states $u_{m\mu}$ for which $E_{m\mu}$ is in ΔE. Since this transition probability is sensibly 0 unless $E_{n\nu}$ is also in ΔE, we conclude that transitions only take place between states of essentially the same energy.

Equation (53·28) now becomes

$$\frac{N_\mu(t) - N_\mu(0)}{t} = -\frac{4\pi^2}{h}\sum_{m \text{ in } \Delta E} N_{m\mu}(0)\sum_\nu J_{\mu\nu}(0) + \frac{4\pi^2}{h}\sum_\nu \sum_{n \text{ in } \Delta E} N_{n\nu}(0)K_{\nu\mu}(0)$$

$$= \frac{4\pi^2}{h}\sum_\nu [N_\nu(0)K_{\nu\mu}(0) - N_\mu(0)J_{\mu\nu}(0)]. \tag{53·29}$$

We denote the quantities $\frac{4\pi^2}{h}K_{\nu\mu}(0)$ and $\frac{4\pi^2}{h}J_{\mu\nu}(0)$ by $\Gamma_{\nu\to\mu}$ and $\Gamma_{\mu\to\nu}$, respectively. They are transition probabilities per unit time, like the Einstein transition probabilities, but have their origin in collisions between molecules rather than in the interaction of matter and radiation. With this notation (53·29) yields the set of differential equations

$$\frac{dN_\mu}{dt} = \sum_\nu (N_\nu\Gamma_{\nu\to\mu} - N_\mu\Gamma_{\mu\to\nu}). \tag{53·30}$$

for the group populations. The particular method which we have used for specifying the groups is evidently not the only possible one.

These transition probabilities are not independent. Let G_μ denote the number of states $u_{m\mu}$ per unit increment in the energy $E_{m\mu}$. Let G_ν denote the corresponding number of states $u_{n\nu}$. These play the parts of statistical weights for the two sets of states. Then

$$G_\mu\Gamma_{\mu\to\nu} = \frac{4\pi^2}{h(\Delta\epsilon)^2}\sum_{E_m \text{ in } \Delta\epsilon}\sum_{E_n \text{ in } \Delta\epsilon}|H_1(m,\mu; n,\nu)|^2$$

$$= G_\nu\Gamma_{\nu\to\mu}. \tag{53·31}$$

Let w_μ, w_ν denote the relative populations per state of the groups μ and ν defined as N_μ/G_μ and N_ν/G_ν, respectively. Equations (53·30) and (53·31) yield

$$\frac{dw_\mu}{dt} = \sum_\nu \Gamma_{\mu\to\nu}(w_\nu - w_\mu). \tag{53·32}$$

From this equation it is clear that groups with more than average population per state must lose population to those with less than average population. *Thus we have a tendency toward an ultimate statistical equilibrium in which the average population per state of each group is equal to that of every other group.* If we reckon the various subjective states $u_{m\mu}$ in any one group as experimentally indistinguishable, we may say that the assemblage as a whole makes an asymptotic approach to a statistical state which is experimentally indistinguishable from one in which all subjective states of the same energy are equally populated. We are accordingly at liberty to use as a first approximation model thermostat assemblage one with a constant diagonal statistical matrix ϱ' whose nonvanishing elements are restricted to a narrow energy range ΔE and have values dependent only on their energies. Allowing ΔE to shrink to zero, thus permitting only *one* energy value, we obtain the model used by Fowler.[1]

53f. The Gibbs Canonical Assemblage for Systems of the Most General Type.—The weakness of the foregoing derivation of (53·32) is that it is based on the use of first-order perturbation approximations for the calculation of increments in w_μ during time intervals which must be large compared with $h/D\epsilon$. The procedure should be satisfactory provided that the fractional change in w_μ for the time interval $h/D\epsilon$ is itself small compared with unity. The most difficult conditions to meet are those in which dw_μ/dt is a maximum, or in which all the w_ν's in (53·32) are set equal to zero. The indicated fractional initial

decrease in w_μ in the time $h/D\epsilon$ is then approximately $\dfrac{h}{D\epsilon}\displaystyle\sum_\nu \Gamma_{\mu\to\nu}$. But

$\displaystyle\sum_\nu \Gamma_{\mu\to\nu}$ is the average total transition probability from states of the μ group to all other groups per unit of time. To estimate its value we note that classically every collision between two molecules would throw the system out of its initial state into another state. Hence $\displaystyle\sum_\nu \Gamma_{\mu\to\nu}$ can be identified with the total number of collisions per unit time in the sample of gas.

If l is the mean free path, v is the average velocity and N the total number of molecules in the sample of gas which constitutes the system, the condition for the validity of (53·32) with all w_ν's set equal to zero is

$$\frac{h}{D\epsilon}\sum_\nu \Gamma_{\mu\to\nu} = \frac{h}{D\epsilon}\frac{Nv}{2l} \ll 1.$$

Taking $D\epsilon$ equal to the average translational kinetic energy per molecule, as previously suggested, we can reduce the above inequality to

[1] *Loc. cit.*

$$\frac{N\lambda}{l} \ll 1$$

where λ is the average de Broglie wave length.

In the asymptotic case in which the volume of the gas is made to approach infinity, while N is held constant, the required conditions can be met and hence we can be assured that (53·32) holds for all w's in the case of an *absolutely* perfect gas. We must apply it with circumspection to actual samples of gas, however, for under standard conditions the value of $N\lambda/l$ for one cubic centimeter of oxygen is of the order of 10^{16}. In the case of initial states close to statistical equilibrium conditions are more favorable, but it is evident that (53·32) does not afford an appropriate basis for the calculation of the approach to statistical equilibrium under ordinary conditions.[1]

In view of this conclusion special interest attaches to the problem of generalizing our discussion of thermostat assemblages to include systems which are not ideal gases. This can be done by means of a reinterpretation of the analysis of Sec. 53c in which we identify the "molecule" of that discussion with a complete macroscopic sample of matter, say S, of any type, gaseous or otherwise, and the "system" with a collection $\mathfrak{S}$ of identical samples $S_1, S_2, \cdots$, each in its own container, but with weak interactions between each sample and its neighbors through the separating walls. We identify H_0 with the sum of the exact Hamiltonians of the separated samples and H_1 with the interaction energy. H_1 can be taken as small as desired, since it is an artificial element in the picture and not an intrinsic property of the sample.

Let us now apply the above analysis to a Gibbsian superassemblage A of identical independent systems $\mathfrak{S}_1, \mathfrak{S}_2, \cdots$. By making H_1 sufficiently small we can avoid the time-interval difficulty and obtain the result that a satisfactory thermostat model for A is one for which all diagonal elements of ϱ' vanish except those associated with a single energy level of H_0, say E_n, and whose nonvanishing diagonal elements are all equal. This is the Fowler model.

Following the method of Darwin and Fowler it is then possible to show that if ϵ_r denotes an energy level of the individual system S the statistical average value of the number of systems S in the level ϵ_r for the model thermostat assemblage A is

$$\bar{a}_r = \frac{e^{-\frac{\epsilon_r}{kT}}}{\sum_n e^{-\frac{\epsilon_n}{kT}}}. \tag{53·33}$$

[1] This should not be surprising since the well-known extreme improbability of appreciable departures from statistical equilibrium implies that when such departures do occur they must be followed by a return to equilibrium which takes place in a very small interval of time.

The proof is given by Fowler,[1] and need not be reproduced here. The number $\bar{a}_r$ is, however, by definition a diagonal element of the statistical matrix ρ for an assemblage of systems S taken from a thermostat at temperature T. The assemblage defined by (53·33) is called a *Gibbs canonical assemblage*. This assemblage is uniquely determined and free from the arbitrariness characteristic of other model thermostat assemblages. It has the property that two assemblages of this kind are not disturbed if interactions are allowed between each member of one assemblage and a corresponding member of the other.

Finally, the canonical assemblage is the only model thermostat assemblage which can be used equally well for microscopic and macroscopic systems. Thus the canonical assemblage may be regarded as the fundamentally correct one even though for convenience in the discussion of macroscopic systems the Fowler model is often preferable.

Equation (53·33) specifies the diagonal matrix ρ for a thermostat assemblage of systems S with nondegenerate energy levels. If S is composed of many like molecules with small interaction energy, one readily verifies that the same formula can be used for the diagonal elements of the corresponding matrix ρ' if we identify r with the general ordinal number for a sequence of orthonormal eigenfunctions of the approximate Hamiltonian and ϵ_r with the corresponding energy.

54. THE ABSORPTION AND EMISSION OF RADIATION: PERTURBATION OF AN ATOMIC SYSTEM BY A CLASSICAL RADIATION FIELD

54a. The Einstein Derivation of the Planck Radiation Formula.— Before taking up the quantum-mechanical theory of the interaction of matter and radiation we pause to give a brief account of Einstein's early "derivation" of the Planck black-body radiation formula.[2] This derivation was in fact a fitting together of the Bohr theory, statistical mechanics, and the radiation formula with the aid of suitable auxiliary hypotheses regarding the frequency of the radiative processes.

Einstein assumed the Bohr postulates **a** and **b** of Sec. 46a and took over from the quantum statistical mechanics of the day (1917) the following formula for the average number of atoms (or molecules) N_n in the nth energy level of the Hamiltonian for the internal coordinates in the case of a thermostat assemblage of samples of gas at the temperature T:

$$N_n = \frac{N g_n^* e^{-\frac{E_n}{kt}}}{\Sigma g_m e^{-\frac{E_m}{kt}}}. \tag{54·1}$$

This formula is readily derived from (53·33). g_n denotes the statistical weight of the energy level, usually identified with the number of different

[1] *Loc. cit.*, §2·32, see footnote 1, p. 439.
[2] A. EINSTEIN, *loc. cit.*, footnote 2, p. 378.

sets of quantum numbers consistent (under a given scheme of quantization) with the energy E_n. From our present point of view g_n is to be identified with the number of linearly independent physically admissible eigenfunctions associated with E_n. If the energy levels are nondegenerate, or if the index n refers to a uniquely defined stationary state with a definite complete set of quantum numbers, the statistical weights are equal to unity.

The "jumps" from one energy level to another which Bohr assumed to accompany the spontaneous emission of radiation are analogous to the spontaneous disintegrations of radioactivity. Hence Einstein supposed that the radiative transitions of free atoms are governed by a law of probability similar to that postulated in the elementary theory of radioactivity. Specifically, he assumed that the probability that an atom in the energy level $E_{n'}$ will have a transition to any definite lower energy level $E_{n''}$ in the time dt is $A_{n' \to n''} dt$, where $A_{n' \to n''}$ is a constant characteristic of the pair of levels called the (Einstein) *transition probability* for the type of jump under consideration. If at any time there are $N_{n'}$ atoms in the level $E_{n'}$, the number of such transitions per second is $N_{n'} A_{n' \to n''}$. Neglecting the Doppler effect due to the translational motions of the atoms, these transitions yield an emission of $N_{n'} A_{n' \to n''} h \nu_{n'n''}$ ergs per second of monochromatic radiation of frequency $\nu_{n'n''} = (E_{n'} - E_{n''})/h$.

In order to take into account the absorption as well as the emission of radiation, the theory postulated a corresponding transition probability for atoms jumping upward from the energy level $E_{n''}$ to the level $E_{n'}$ in a radiation field. In accordance with experiment the probability of this second type of transitions was assumed to be proportional to the energy density of the radiation field per unit frequency interval at the frequency $\nu_{n'n''}$. Let $\rho(\nu_{n'n''})$ denote this quantity. Einstein assumed that the number of transitions of this second type per second is $N_{n''} B_{n'' \to n'} \rho(\nu_{n'n''})$, where $B_{n'' \to n'}$ is a second transition probability.

Still a third type of transition, *viz.*, emission from the energy level $E_{n'}$ stimulated by the radiation field and proportional to $\rho(\nu_{n'n''})$ was needed in order to account for the desired Planck formula. The existence of such stimulated emission was foreshadowed by the classical theory of the interaction between an oscillating system of charges and a classical radiation field.[1] The probable number of such emission processes per second was set equal to $N_{n'} B_{n' \to n''} \rho(\nu_{n'n''})$, where $B_{n' \to n''}$ is a third transition probability.

In thermal equilibrium, emission and absorption must balance so that

$$N_{n'}[A_{n' \to n''} + B_{n' \to n''} \rho(\nu_{n'n''})] = N_{n''} B_{n'' \to n'} \rho(\nu_{n'n''}).$$

Inserting the values of $N_{n'}$, $N_{n''}$ and simplifying, we obtain

[1] *Cf., e.g.,* J. H. VAN VLECK, *Phys. Rev.* **24**, 330 (1924).

$$g_{n'}[A_{n'\to n''} + B_{n'\to n''}\rho(\nu_{n'n''})] = g_{n''}e^{\frac{h\nu_{n'n''}}{kT}}B_{n''\to n'}\rho(\nu_{n'n''}).$$

Hence

$$\rho(\nu_{n'n''}) = \frac{A_{n'\to n''}}{B_{n'\to n''}\left(\dfrac{g_{n''}B_{n''\to n'}}{g_{n'}B_{n'\to n''}}e^{\frac{h\nu_{n'n''}}{kT}} - 1\right)}. \tag{54·2}$$

In order to deduce the Planck formula

$$\rho(\nu) = \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{\frac{h\nu}{kT}} - 1}, \tag{54·3}$$

Einstein had only to assume the relationships

$$g_{n''}B_{n''\to n'} = g_{n'}B_{n'\to n''}, \tag{54·4}$$

$$A_{n'\to n''} = \frac{8\pi h\nu_{n'n''}{}^3}{c^3}B_{n'\to n''}. \tag{54·5}$$

With the aid of these formulas it was possible to derive the values of $B_{n'\to n''}$ and $B_{n''\to n'}$ from the corresponding value of $A_{n'\to n''}$ when the last mentioned transition probability had been worked out from the Bohr correspondence principle. Although thermal equilibrium must be assumed in order to obtain the radiation formula, the postulate that atoms jump from the upper level to the lower at a net rate

$$N_{n'}[A_{n'\to n''} + B_{n'\to n''}\rho(\nu_{n'n''})] - N_{n''}B_{n''\to n'}\rho(\nu_{n'\,n''}) \tag{54·6}$$

is assumed to hold outside thermal equilibrium for any "natural" (chaotic) radiation field.

54b. Elementary Approaches to the Quantum Theory of the Einstein Transition Probabilities.—As noted in Sec. 53*d* Bohr's basic assumption that every atom or molecule spends practically all its time in one or another of the stationary definite energy states with intermittent discontinuous jumps from one to the other does not hold in quantum mechanics. Only in a very special class of subjective states is an atom assigned to one definite energy level. Nevertheless it is possible to give a satisfactory reinterpretation of the concepts which enter his formulas and to maintain that the formulas are still correct. ʼA complete quantum theory of the emission and absorption of radiation derives the Einstein formulas and the values of the transition probabilities which appear in them. Such a theory was first given by Dirac[1] following an argument similar to that of Sec. 53*e*, but with terms in the Hamiltonian for the energy of the radiation field and the mutual energy of matter and field. To give an account of the Dirac theory would lead us beyond the scope of this book. Hence we shall content ourselves with the lesser

---

[1] P. A. M. DIRAC, *loc. cit.;* E. FERMI, *Rev. Mod. Phys.* **4,** 87 (1932); G. BREIT, *Rev. Mod. Phys.* **4,** 504 (1932).

problem of setting up formulas for the basic constants $A_{n'\to n''}$, $B_{n'\to n''}$, $B_{n''\to n'}$.

This problem can be dealt with in a variety of ways involving different degrees of sophistication. Schrödinger's original "hydrodynamical" method[1] treated the atom as a continuous charge distribution to which the classical radiation formulas could be directly applied. The charge density $\rho(x,y,z,t)$ was computed from the wave function $\Psi(x_1, \cdots, z_f, t)$ as we should now compute the instantaneous statistical average charge density. The procedure is out of harmony with the present interpretation of the formalism of wave mechanics and gives results which are only partly right, since the computed rate of spontaneous emission depends on the "population" of the lower energy levels into which the atoms drop as they emit. In fact the computed rate of spontaneous emission falls to zero if all the atoms of the assemblage described by a wave function Ψ are confined to a single-energy level.

In the Heisenberg matrix theory the correct formulas for the Einstein transition probabilities are obtained directly from the Bohr correspondence principle, *i.e.*, as *ad hoc* postulates based on classical analogy. Thus the rate of spontaneous emission per atom due to transitions from one nondegenerate state to another given by (46·8) is to be identified with the quantity $h\nu_{n'n''}A_{n'\to n''}$. The transition probability is accordingly

$$A_{n'\to n''} = \frac{64\pi^4\nu_{n'n''}^3}{3hc^3}|\vec{D}(n',n'')|^2 \tag{54·7}$$

where $\vec{D}$ is the Heisenberg matrix associated with the electric moment of the atom, $\vec{D} = \sum_k e_k\vec{r}_k$. In order to work out the corresponding Einstein transition probabilities for degenerate energy levels we proceed as follows. We assume that we have to do with an assemblage in which the population $N_{n'm'}$, of a substate of $E_{n'}$, defined by the wave function $\psi_{n'm'}$, is independent of m'. (This is in accordance with the usual rules of elementary quantum statistical mechanics for thermal equilibrium.) Then the probability that an arbitrary system in the level $E_{n'}$ will make a transition to $E_{n''}$ in unit time is the average of the corresponding probabilities for the various initial substates. The probability for the substate defined by $\psi_{n'm'}$ is $\frac{64\pi^4\nu_{n'n''}^3}{3hc^3}\sum_{m''}|\vec{D}(n',m';\ n'',m'')|^2$. Taking the average over the $g_{n'}$ substates of the upper energy level we obtain

$$A_{r'\to n''} = \frac{64\pi^4\nu_{n'n''}^3}{3hc^3g_{n'}}\sum_{m'}\sum_{m''}|\vec{D}(n',m';\ n'',m'')|^2. \tag{54·8}$$

[1] E. SCHRÖDINGER, *Ann. d. Physik* **81**, 134 (1926). *Cf.* also A. Sommerfeld, *A.S.W.E.*, p. 56, and Condon and Morse, *Q.M.*, p. 90.

From Eqs. (54·5) and (54·4) we learn that the corresponding values of the other two Einstein transition probabilities are

$$B_{n' \to n''} = \frac{8\pi^3}{3h^2 g_{n'}} \sum_{m'} \sum_{m''} |\vec{D}(n',m'; n'',m'')|^2, \qquad (54·9)$$

$$B_{n'' \to n'} = \frac{8\pi^3}{3h^2 g_{n''}} \sum_{m'} \sum_{m''} |\vec{D}(n',m'; n'',m'')|^2. \qquad (54·10)$$

These results are in harmony with experiment and with the conclusions to be obtained from more complete theories.

A third method of attack on the problem of emission and absorption, independently devised by Dirac, Born, and Slater, is to deal with absorption and stimulated emission by a direct application of perturbation theory to the problem of the interaction of a quantum-mechanical system with a classical radiation field.[1] Using this procedure we obtain the formulas (54·9) and (54·10) as a consequence of principles already laid down without recourse to analogy. Unfortunately it is impossible to give a direct quantum-mechanical treatment of the problem of spontaneous emission except by applying the same quantum principles to the radiation field which we have developed for atomic systems.[2] Hence the procedure under consideration has to lean on the Einstein theory for a derivation of $A_{n' \to n''}$ from $B_{n' \to n''}$.

The remainder of Sec. 54 is devoted to an exposition of the method of Dirac, Born, and Slater.

54c. The Perturbing Hamiltonian for a Classical Radiation Field.— The Hamiltonian operator for an atomic system interacting with a classical radiation field with the scalar potential $\Phi(x,y,z)$ and the vector potential $\vec{\alpha}(x,y,z)$ was derived in Sec. 7, p. 28. The reduced form of this Hamiltonian [cf. Eqs. (7·11) and (49·2)] is

$$H = -\sum_{j=1}^{n} \frac{h^2}{8\pi^2 \mu_j} \left[\nabla_j^2 - \frac{4\pi i e_j}{hc} \vec{\alpha}^{(j)} \cdot \nabla_j + \frac{2\pi i e_j}{hc^2} \frac{\partial \Phi^{(j)}}{\partial t} \right] + V', \qquad (54·11)$$

where

$$\begin{aligned}
\vec{\alpha}^{(j)} &= \vec{\alpha}(x_j, y_j, z_j, t), \qquad \Phi^{(j)} = \Phi(x_j, y_j, z_j, t), \\
V' &= V + \sum_{j=1}^{n} \left[e_j \Phi_j + \frac{e_j^2}{2c^2 \mu_j} |\vec{\alpha}^{(j)}|^2 \right].
\end{aligned} \qquad (54·12)$$

---

[1] Cf. P. A. M. Dirac, *Proc. Roy. Soc.* **A112**, 673ff (1926); M. Born, *Zeits. f. Physik* **40**, 167 (1926); J. C. Slater, *Proc. Nat. Acad. Sci.* **13**, 7 (1927).

[2] Cf. footnote 1, p. 450.

It can be easily verified that this Hamiltonian is Hermitian with respect to class D and the domain of all Cartesian-coordinate space. This was almost proved for the three-dimensional case in Sec. 8, where we showed that

$$\frac{\partial}{\partial t}\int\int\int \Psi\Psi^* dx dy dz = 0,$$

provided that Ψ is a suitably restricted solution of the Schrödinger equation

$$H\Psi = -(h/2\pi i)\partial\Psi/\partial t,$$

with H defined by Eq. (54·11), n being set equal to unity. The reader will readily verify that the conditions for the theorem there proved are fulfilled if Ψ is of class D. To prove the Hermitian character of H, we replace Ψ by $\Psi_1 + \Psi_2$ and $\Psi_1 + i\Psi_2$, in turn, where Ψ_1 and Ψ_2 are any two class D solutions of the differential equation. It follows that

$$\frac{\partial}{\partial t}\int\int\int_\infty (\Psi_1\Psi_2^* + \Psi_2\Psi_1^*) dx dy dz = 0$$

and

$$\frac{\partial}{\partial t}\int\int\int_\infty (\Psi_1\Psi_2^* - \Psi_2\Psi_1^*) dx dy dz = 0.$$

Hence

$$\frac{\partial}{\partial t}\int\int\int_\infty \Psi_1\Psi_2^* dx dy dz = -\frac{2\pi i}{h}\int\int\int (\Psi_2^* H\Psi_1 - \Psi_1 H\Psi_2^*) dx dy dz = 0.$$

Since we can choose any class D function as a possible instantaneous form for a solution of the Schrödinger equation, it follows that

$$(H\Psi_1,\Psi_2) = (\Psi_1,H\Psi_2)$$

holds for any pair of class D functions, or that H is Hermitian with respect to class D. The extension of the theorem to the case of a many-particle problem is left to the reader.

Let H_0 denote the Hamiltonian for the atomic system in the absence of a radiation field. As in Sec. 53 we assume an arbitrary subjective state $\Psi(x,t)$ and an expansion in terms of a normal orthogonal set of discrete eigenfunctions $\psi_n(x)$ of H_0. Thus

$$\Psi(x,t) = \sum_n \xi_n(t)\psi_n(x)e^{-\frac{2\pi i E_n t}{h}} = \sum_n \xi_n(t)\Psi_n(x,t). \qquad (54·13)$$

We shall compute the variation in the coefficients ξ_n with time due to the radiation field.

Our purpose is to determine the rate of change in the relative probability of the different states ψ_n as affected by interaction with a "natural" radiation field of given frequency distribution (as given by experimental spectrum analysis). In making such a computation we calculate the perturbation in a time interval short enough to allow us to treat the radiation energy density per unit frequency interval as constant and the changes in the ξ_n's as infinitesimals. It is important to note, however, that the time interval under consideration, while small from a macroscopic

point of view, is best made large from a microscopic point of view. Otherwise we should get into difficulty in relating the convenient mathematical description of the field with the empirical spectrum analysis (*cf.* pp. 457 and 458).

Let us deal first with the elementary case of a radiation field consisting solely of plane waves moving in the direction of the x axis and polarized in the x,z plane, *i.e.*, with the electric vector parallel to the y axis. In the case of a plane wave system the scalar potential Φ can be set equal to zero, in which case the vector potential is purely transverse and parallel to the electric vector. Let the interval during which the perturbing field acts at the origin be $0 < t < \theta$. As we are concerned only with the values of the vector potential during this interval we can represent the potential by a Fourier integral. Thus

$$\mathfrak{a}_y = \lambda f\left(t - \frac{x}{c}\right) = \lambda c \int_{-\infty}^{+\infty} \frac{\phi_y(\nu)}{2\pi i \nu} e^{2\pi i \nu \left(t - \frac{x}{c}\right)} d\nu, \qquad \mathfrak{a}_x = \mathfrak{a}_z = 0, \quad (54 \cdot 14)$$

$$\mathcal{E}_y = -\frac{1}{c} \frac{\partial \mathfrak{a}_y}{\partial t} = -\lambda \int_{-\infty}^{+\infty} \phi_y(\nu) e^{2\pi i \nu \left(t - \frac{x}{c}\right)} d\nu, \qquad \mathcal{E}_x = \mathcal{E}_z = 0. \quad (54 \cdot 15)$$

Here λ is a real parameter which determines the amplitude of the wave system and which will be identified with the parameter λ of the general perturbation theory of Sec. 53. As $f(t - x/c)$ is real, the amplitude function for the electric force $\phi_y(\nu)$ is subject to the relation

$$\phi_y(\nu) = \phi_y{}^*(-\nu).$$

Dropping the scalar potential Φ from (54·11) and neglecting the terms $\sum_j e_j |\vec{\mathfrak{a}}^{(j)}|^2 / 2c^2 \mu_j$ in V', we reduce the Hamiltonian to the standard form $H_0 + \lambda H_1$, with

$$\lambda H_1 = -\frac{h}{2\pi i c} \sum_j \frac{e_j}{\mu_j} \left(\mathfrak{a}_x{}^{(j)} \frac{\partial}{\partial x_j} + \mathfrak{a}_y{}^{(j)} \frac{\partial}{\partial y_j} + \mathfrak{a}_z{}^{(j)} \frac{\partial}{\partial z_j} \right). \quad (54 \cdot 16)$$

Equation (54·14) gives the further reduction

$$H_1 = -\frac{h}{2\pi i c} \sum_j \frac{e_j}{\mu_j} f\left(t - \frac{x_j}{c}\right) \frac{\partial}{\partial y_j}. \quad (54 \cdot 17)$$

54d. The Born Transition Probability.—The first step in carrying through the perturbation scheme of Sec. 53 is to evaluate the elements of the Heisenberg matrix $\mathbf{H}_1^{(t)}$. A typical element has the form

$$H_1^{(t)}(n,m) = \int_{\infty} \cdots \int \Psi_n{}^* H_1 \Psi_m d\tau. \quad (54 \cdot 18)$$

If the wave lengths which contribute appreciably to the field are all large

compared with the atomic dimensions, we can treat the vector $\vec{a}$ in first approximation as constant over the neighborhood of the atom and can replace $f\left(t - \frac{x_j}{c}\right)$ by the value of $f\left(t - \frac{x}{c}\right)$ at the origin, which we take to be at or near the center of the system. This approximation is equivalent to neglecting the absorption due to the magnetic dipole moment and the electric quadrupole moment in comparison with that due to the electric dipole moment (*cf.* Sec. 54*g*). Then

$$H_1^{(t)}(n,m) = -\frac{f(t)}{c}\sum_j \frac{e_j}{\mu_j} \int \cdots \int \Psi_n^* \frac{h}{2\pi i} \frac{\partial}{\partial y_j} \Psi_m d\tau$$

$$= -\frac{f(t)}{c}\sum_j \frac{e_j}{\mu_j} p_{y_j}^{(t)}(n,m). \qquad (54\cdot19)$$

Here $\mathbf{p}_y^{(t)}$ is the Heisenberg matrix of p_y for the unperturbed system. It follows from the first of the Hamiltonian equations (45·4) that

$$\mathbf{p}_{y_j}^{(t)} = \mu_j \frac{d}{dt} \mathbf{y}_j^{(t)}. \qquad (54\cdot20)$$

Hence

$$\mathbf{H}_1^{(t)} = -\frac{f(t)}{c} \frac{d}{dt}\sum_j e_j \mathbf{y}_j^{(t)} = -\frac{f(t)}{c} \frac{d}{dt}\mathbf{D}_y^{(t)}, \qquad (54\cdot21)$$

where $\mathbf{D}_y^{(t)}$ is the Heisenberg matrix of the y component of the electric moment. Carrying out the indicated differentiation with respect to the time, we obtain

$$H_1^{(t)}(n,m) = \frac{2\pi i}{c} \frac{(E_m - E_n)}{h} f(t) D_y^{(t)}(n,m). \qquad (54\cdot22)$$

It follows from this equation that the elements of $\mathbf{H}_1^{(t)}$ between states of the same energy are all zero. With the aid of Eqs. (53·11) and (53·17) we infer that dipole radiation produces no transitions between different states of the same energy.

The first-order perturbations of the wave functions are now determined by Eqs. (53·9), (53·12), and (53·13):

$$T_1(\theta)^{(1)}\{n,m\} = \frac{4\pi^2}{h^2 c}\int_0^\theta (E_m - E_n)f(t')D_y^{(t')}(n,m)dt'. \qquad (54\cdot23)$$

Let

$$\nu_{nm} = \frac{E_n - E_m}{h}; \qquad D_y(n,m) = e^{-2\pi i\nu_{nm}t}D_y^{(t)}(n,m) = \int_\infty \psi_n^* D_y \psi_m d\tau.$$

$$(54\cdot24)$$

The matrix $\mathbf{D}_y$ is a Schrödinger matrix based on the arbitrary ortho-normal set of eigenfunctions of H_0 introduced in (54·13). It follows from the Fourier integral theorem (Sec. 9a) and Eq. (54·14) that

$$\frac{c\phi_y(\nu)}{2\pi i\nu} = \int_{-\infty}^{+\infty} f(t)e^{-2\pi i\nu t}dt. \tag{54·25}$$

As we are not interested in $f(t)$ outside the interval $0 < t < \theta$, we can assume that it vanishes outside the interval in question and can modify the limits of integration accordingly. Substituting ν_{nm} for ν, we obtain

$$\frac{c\phi_y(\nu_{nm})}{2\pi i\nu_{nm}} = \int_0^\theta f(t)e^{-2\pi i\nu_{nm}t}dt. \tag{54·26}$$

A combination of this result with Eq. (54·23) yields

$$F^{(1)}(n,m) = T_1(\theta)^{(1)}\{n,m\} = -\frac{2\pi i}{h}\phi_y(\nu_{nm})D_y(n,m), \tag{54·27}$$

where $\mathbf{F}^{(1)}$ is simply a convenient abbreviation for $\mathbf{T}_1(\theta)^{(1)}$.

The Born transition probability, giving the chance that a system initially in the state ψ_n will be found after the time θ in the state ψ_m, is now given in first approximation by (53·17). If $n \neq m$, we find to terms in λ^3,

$$\Phi_{n\to m} = \frac{4\pi^2}{h^2}\lambda^2|\phi_y(\nu_{nm})|^2|D_y(n,m)|^2. \tag{54·28}$$

The chance that a system initially in the state ψ_n will be found in the same state at the end of the perturbation is

$$\Phi_{n\to n} = 1 - \frac{4\pi^2}{h^2}\lambda^2\sum_{m\neq n}|\phi_y(\nu_{nm})|^2|D_y(n,m)|^2 \tag{54·29}$$

to terms in λ^3.

So far we have had to do with the effect of a single plane-polarized plane-wave train on the motion and energy of an isolated atomic system. The important physical problem, however, is that of the perturbation produced by a "natural," or chaotic, radiation field. Such a field can be analyzed into a superposition of plane waves moving in all directions with a random distribution of phases. It suffices for our present purpose, however, to analyze the components of $\vec{\alpha}$ at the origin into Fourier integrals. Thus we write

$$\left.\begin{aligned}
\alpha_x &= \lambda c\int_{-\infty}^{+\infty}\frac{\phi_x(\nu)}{2\pi i\nu}e^{2\pi i\nu t}d\nu, & \mathcal{E}_x &= -\lambda\int_{-\infty}^{+\infty}\phi_x(\nu)e^{2\pi i\nu t}d\nu,\\
\alpha_y &= \lambda c\int_{-\infty}^{+\infty}\frac{\phi_y(\nu)}{2\pi i\nu}e^{2\pi i\nu t}d\nu, & \mathcal{E}_y &= -\lambda\int_{-\infty}^{+\infty}\phi_y(\nu)e^{2\pi i\nu t}d\nu,\\
\alpha_z &= \lambda c\int_{-\infty}^{+\infty}\frac{\phi_z(\nu)}{2\pi i\nu}e^{2\pi i\nu t}d\nu, & \mathcal{E}_z &= -\lambda\int_{-\infty}^{+\infty}\phi_z(\nu)e^{2\pi i\nu t}d\nu.
\end{aligned}\right\} \tag{54·30}$$

In evaluating ϕ_x, ϕ_y, ϕ_z we again assume that $\vec{\mathcal{E}}$ and $\vec{\alpha}$ vanish at the origin outside the interval $0 < t < \theta$.

The contributions of the three orthogonal components of $\vec{\alpha}$ to the matrix $\mathbf{F}^{(1)}$ are readily seen to be independent. Hence

$$F^{(1)}\{n,m\} = -\frac{2\pi i \lambda}{h}[\phi_x(\nu_{nm})D_x(n,m) + \phi_y(\nu_{nm})D_y(n,m)$$
$$+ \phi_z(\nu_{nm})D_z(n,m)]. \quad (54\cdot31)$$

In this more general case "nondiagonal elements" of the Born transition probability are obtained by multiplying the right-hand member of (54·31) by its complex conjugate. Some of the terms are essentially positive, but the phases of those involving products of field-strength components in different directions, such as $\phi_x(\nu_{nm})\phi_y^*(\nu_{nm})$ can have any value. Since we never know the phases of the components of a radiation field, we are interested in the mean value of $\Phi_{n \to m}$ for all possible phases. In such an average the above-mentioned cross terms drop out. If n is not equal to m, we have, in first approximation,

$$\overline{\Phi_{n \to m}} = \frac{4\pi^2\lambda^2}{h^2}\{\overline{|\phi_x(\nu_{nm})|^2}|D_x(n,m)|^2 + \overline{|\phi_y(\nu_{nm})|^2}|D_y(n,m)|^2 +$$
$$\overline{|\phi_z(\nu_{nm})|^2}|D_z(n,m)|^2\}. \quad (54\cdot32)$$

According to Plancherel's theorem

$$\int_0^\theta (\mathcal{E}_x{}^2 + \mathcal{E}_y{}^2 + \mathcal{E}_z{}^2)dt = \lambda^2 \int_{-\infty}^{+\infty}\{|\phi_x|^2 + |\phi_y|^2 + |\phi_z|^2\}d\nu.$$

Hence the average energy density of the field for the period of the perturbation $0 < t < \theta$ is

$$\frac{\overline{\mathcal{E}^2}}{4\pi} = \int_0^\infty w(\nu)d\nu, \quad (54\cdot33)$$

where

$$w(\nu) = \frac{\lambda^2}{2\pi\theta}\{|\phi_x(\nu)|^2 + |\phi_y(\nu)|^2 + |\phi_z(\nu)|^2\}. \quad (54\cdot34)$$

The energy density per unit frequency interval $\rho(\nu)$ which appears in Eqs. (54·2) and (54·3) is the average value of $w(\nu)$ for a physically infinitesimal frequency interval wide enough to smooth out the violent microscopic fluctuations characteristic of the latter, but small compared with the intervals in which there is an appreciable macroscopic experimental intensity variation.[1] In the case of a radiation field which is

---

[1] This statement involves the tacit assumption that $1/\theta$ is small compared with ν_{nm}. Otherwise $\overline{w(\nu)}$ would not agree with the energy density observed with a normal photographic exposure.

macroscopically isotropic an averaging over such a physically infinitesimal frequency interval yields

$$\overline{|\phi_x(\nu)|^2} = \overline{|\phi_y(\nu)|^2} = \overline{|\phi_z(\nu)|^2} = \frac{2\pi\theta}{3\lambda^2}\rho(\nu). \tag{54.35}$$

Hence

$$\overline{\Phi_{n\to m}} = \frac{8\pi^3\theta}{3h^2}\rho(\nu_{nm})\{|D_x(n,m)|^2 + |D_y(n,m)|^2 + |D_z(n,m)|^2\} \tag{54.36}$$

if $n \neq m$.

54e. The Einstein Transition Probabilities.—It is necessary only to divide the expression (54·36) for the Born transition probability by the time interval θ in order to obtain a formal derivation of the Einstein transition probability $B_{n\to m}$ for transitions between uniquely defined (nondegenerate) states. The formulas (54·9) and (54·10) for the degenerate case are then obtainable by the procedure previously employed in deriving (54·8) from (54·7). It will be observed that our method of approach gives the correct values for the transition probabilities associated with induced emission as well as with absorption.

The above procedure must not be accepted as satisfactory without critical discussion, however, since the application of the Einstein transition probabilities to thermal equilibrium problems and to the prediction of the intensities of spectrum lines in absorption and emission involves much more complicated situations than that envisaged in the derivation of (54·36).

In applying the Einstein theory of Sec. 54a to a sample of a pure gas it is customary to assume that the radiation from each molecule is incoherent with that from the others so that the total radiation emitted and absorbed per unit time is the same as for a Gibbsian assemblage of N independent molecules in an appropriately defined mixed subjective state. We have accordingly to compute the time derivatives of the diagonal elements of the statistical matrix ρ' for a Gibbsian assemblage of individual molecules in a chaotic mixed state. Employing the same procedure as in Sec. 53e we again derive Eqs. (53·26) with N_n interpreted as the number of molecules in the state whose wave function is ψ_n. Equation (53·27) becomes

$$N_m(\theta) - N_m(0) = \frac{4\pi^2\lambda^2}{h^2}\sum_n[N_n(0) - N_m(0)]\{|\phi_x(\nu_{nm})|^2|D_x(n,m)|^2$$
$$+ |\phi_y(\nu_{nm})|^2|D_y(n,m)|^2 + |\phi_z(\nu_{nm})|^2|D_z(n,m)|^2\}. \tag{54.37}$$

If the radiation field is isotropic, (54·37) reduces to

$$N_m(\theta) - N_m(0) = \frac{8\pi^3\theta}{3h^2}\sum[N_n(0) - N_m(0)]\rho(\nu_{nm})|\vec{D}(n,m)|^2. \tag{54.38}$$

Even when the field is not isotropic, it is possible to effect a reduction of (54·37) similar to (54·38) by means of the theorem that the sums

$$\sum_n \sum_m |D_x(n,m)|^2, \quad \sum_n \sum_m |D_y(n,m)|^2, \quad \sum_n \sum_m |D_z(n,m)|^2 \text{ are equal when each is}$$

extended over all pairs of states n,m belonging to the same energy level. This theorem we now proceed to prove.

It is convenient to return to the double indexing system of Eqs. (54·8), (54·9), and (54·10). The typical matrix element of an arbitrary dynamical variable α is then given by

$$\alpha(n',m'; n'',m'') = \int \psi^*_{n'm'} \alpha \psi_{n''m''} d\tau.$$

Let $^{(n',n'')}\alpha$ denote the matrix of $g_{n'}$ rows and $g_{n''}$ columns derived from $\alpha(n',m'; n'',m'')$ by allowing m' and m'' to range through all values compatible with specific values of n' and n'', respectively. If we multiply $^{(n',n'')}\alpha$ by its adjoint $^{(n',n'')}\alpha^\dagger$, we obtain a Hermitian square matrix of order $g_{n'}$ or $g_{n''}$ according as $^{(n',n'')}\alpha$, or $^{(n',n'')}\alpha^\dagger$, is the antecedent factor. In either case the diagonal sum of the product is

$$\sum_{m'} \sum_{m''} |\alpha(n'm'; n''m'')|^2 = \text{Spur} \left({}^{(n',n'')}\alpha \, {}^{(n',n'')}\alpha^\dagger \right).$$

By the general theorem of p. 358, Sec. 44*b*, this quantity is invariant of a unitary transformation.

Let us apply the above lemma to the partial matrices $^{(n',n'')}\mathbf{D}_x$, $^{(n',n'')}\mathbf{D}_y$, $^{(n',n'')}\mathbf{D}_z$ of the components of the electric moment of the molecule or atom under consideration. We assume that the unperturbed Hamiltonian H_0 is symmetrical with respect to the three coordinate axes so that all three components of the angular momentum are conserved. The basic wave functions of the matrix representation can then be chosen to be simultaneous eigenfunctions of H_0, $\mathfrak{L}^2$, $\mathfrak{L}_z$. In this case the matrices $^{(n',n'')}\mathbf{D}_x$, $^{(n',n'')}\mathbf{D}_y$, $^{(n',n'')}\mathbf{D}_z$ can be transformed one into the other by suitably chosen canonical transformations. Thus, in order to establish the relation

$$^{(n',n'')}\mathbf{D}_z = \mathbf{U}^{-1} \, {}^{(n',n'')}\mathbf{D}_x \mathbf{U},$$

it is only necessary to choose for $\mathbf{U}$ the matrix of the transformation which replaces simultaneous eigenfunctions of H_0, $\mathfrak{L}^2$, $\mathfrak{L}_z$ by simultaneous eigenfunctions of H_0, $\mathfrak{L}^2$, $\mathfrak{L}_x$. Clearly the matrix representation of D_x in the latter system of coordinate functions is identical with the matrix representation of D_z in the former system. Since the diagonal sums are invariant under such a transformation they must have the same initial values for $^{(n',n'')}\mathbf{D}_x$, $^{(n',n'')}\mathbf{D}_y$, $^{(n',n'')}\mathbf{D}_z$, as was to be proved.

At this point we introduce the double index notation into Eq. (54·37). It becomes

$$N_{n''m''}(\theta) - N_{n''m''}(0) = \frac{4\pi^2\lambda^2}{h^2}\sum_{n'}\sum_{m'}[N_{n'm'}(0) - N_{n''m''}(0)]$$

$$\times \{|\phi_x(\nu_{n'n''})|^2|D_x(n'm';n''m'')|^2 + |\phi_y(\nu_{n'n''})|^2|D_y(n'm';n''m'')|^2$$
$$+ |\phi_z(\nu_{n'n''})|^2|D_z(n'm';n''m'')|^2\}. \quad (54\cdot39)$$

Let N_n denote the average number of molecules in the energy level E_n, i.e., let $N_n = \sum_m N_{nm} = N\sum_m|\xi_{nm}|^2$. Summing (54·39) over the $g_{n''}$ substates of the level $E_{n''}$, we obtain

$$N_{n''}(\theta) - N_{n''}(0) = \frac{4\pi^2\lambda^2}{h^2}\sum_{n'}\sum_{m'}\sum_{m''}[N_{n'm'}(0) - N_{n''m''}(0)]$$

$$\times \{|\phi_x(\nu_{n'n''})|^2|D_x(n'm';n''m'')|^2 + |\phi_y(\nu_{n'n''})|^2|D_y(n'm';n''m'')|^2$$
$$+ |\phi_z(\nu_{n'n''})|^2|D_z(n'm';n''m'')|^2\}.$$

We assume, as in Sec. 54b, that the population N_{nm} of the different substates of any energy level E_n are equal. Hence the quantity

$$[N_{n'm'}(0) - N_{n''m''}(0)]$$

is actually independent of the indices m' and m''. It can be written as

$\left[\dfrac{N_{n'}(0)}{g_{n'}} - \dfrac{N_{n''}(0)}{g_{n''}}\right]$ and taken out from under the signs $\sum\limits_{m'}$, $\sum\limits_{m''}$.

In virtue of the diagonal sum rule the above equation reduces on averaging over a chaotic assemblage of radiation fields to

$$N_{n''}(\theta) - N_{n''}(0) = \frac{8\pi^3\theta}{3h^2}\sum_{n'}\rho(\nu_{n'n''})\left[\frac{N_{n'}(0)}{g_{n'}}\right.$$

$$\left. - \frac{N_{n''}(0)}{g_{n''}}\right]\sum_{m'}\sum_{m''}|\vec{D}(n'm';n''m'')|^2. \quad (54\cdot40)$$

Recalling that the time interval θ is by hypothesis an infinitesimal from the large-scale point of view, we adopt the notation $\Delta N_{n''}/\Delta t$ for the ratio $[N_{n''}(\theta) - N_{n''}(0)]/\theta$. The theoretical expression for the mean rate of increase of $N_{n''}$ in a chaotic radiation field now takes the form

$$\frac{\Delta N_{n''}}{\Delta t} = \sum_{n'}\rho(\nu_{n'n''})[N_{n'}B_{n'\to n''} - N_{n''}B_{n''\to n'}], \quad (54\cdot41)$$

where $B_{n'\to n''}$ and $B_{n''\to n'}$ have the values of Eqs. (54·9) and (54·10) and so are to be identified with the Einstein transition probabilities for induced jumps[1] between the energy levels $E_{n'}$ and $E_{n''}$.

[1] We retain the language of the Bohr theory and speak of "quantum jumps" and "transition probabilities" in spite of the fact that the radiation field produces only continuous changes in the wave functions describing the assemblage under

The complete Einstein law includes the spontaneous transition probability $A_{n'\to n''}$ from the energy level $E_{n'}$ to each *lower* level $E_{n''}$. With the addition of terms of this type the expression for $\Delta N_{n''}/\Delta t$ should be

$$\frac{\Delta N_{n''}}{\Delta t} = \sum_{\text{all } n'} \rho(\nu_{n'n''})[N_{n'}B_{n'\to n''} - N_{n''}B_{n''\to n'}]$$

$$+ \sum_{E_{n'}>E_{n''}} N_{n'}A_{n'\to n''} - \sum_{E_{n'}<E_{n''}} N_{n''}A_{n''\to n'}, \quad (54.42)$$

where $A_{n'\to n''}$ has the value (54·8) in virtue of (54·5). As previously stated, a complete quantum-mechanical derivation of (54·42), independent of the quasi-thermodynamic theory of Einstein, would treat the field, as well as the molecule which interacts with it, on a quantum-mechanical basis.

Our present theory is not only weak in its failure to give an account of spontaneous emission. It treats the radiation field as a known *external* perturbing influence and so really fails to tell us anything about the reaction of the molecule on the field. We must infer from the general law of the conservation of energy, however, that the energy absorbed or emitted by the gas is compensated by energy lost or gained by the radiation field as the case may be. Evidently the phases of the radiation waves spontaneously emitted must be independent of the external radiation-field phases. On the other hand classical theory would suggest, and experiment verifies, that the energy absorbed or emitted as the result of induced transitions is compensated by spherical radiation wavelets *coherent* with the inducing wave system and producing a damping of the primary waves as they pass through the gas. Hence the terms of (54·42) involving $A_{n'\to n''}$ and $A_{n''\to n'}$ are associated with the emissivity of the gas, while the terms in $B_{n'\to n''}$ and $B_{n''\to n'}$ are associated with the absorption coefficient.

In applying our results to the study of the radiation absorbed and emitted we have to remember that the spectrum line associated with any pair of energy levels $E_{n'},E_{n''}$ is not infinitely narrow, as our approximate theory would indicate, but has a finite frequency breadth, say 2η. We accordingly introduce the emissivity per unit volume per unit frequency interval[1] ϵ_ν, and identify the integral of this quantity over the line breadth with the total energy radiated per unit volume as a result of spontaneous transitions from one level to the other. Let the numbers

consideration. Any finite change in a wave function involves a redistribution of probability for the eigenvalues of an observable α which suggests the existence of jumps from one eigenstate to another. Thus the idea of discontinuous transitions gets into the theory in the interpretation of the wave function as a probability amplitude.

[1] *Cf.* M. PLANCK, *Wärmestrahlung*, p. 7, Leipzig, 1933.

$N_{n'}$ and $N_{n''}$ be the populations *per unit volume* of the corresponding levels. Then the integrated emissivity for the line is

$$\epsilon(n',n'') = \int_{\nu_{n'n''}-\eta}^{\nu_{n'n''}+\eta} \epsilon_\nu d\nu = N_{n'}A_{n'\to n''}h\nu_{n'n''}. \qquad (54\cdot43)$$

The energy absorbed per unit volume per unit time in any given frequency range $\nu_1 < \nu < \nu_2$ is equal to $c\int_{\nu_1}^{\nu_2}\rho(\nu)\alpha_\nu d\nu$, where α_ν is the absorption coefficient. The integrated value of the absorption coefficient over the spectrum line is therefore

$$\alpha(n',n'') = \int_{\nu_{n'n''}-\eta}^{\nu_{n'n''}+\eta} \alpha_\nu d\nu = [N_{n'}B_{n'\to n''} - N_{n''}B_{n''\to n'}]\frac{h\nu_{n'n''}}{c}. \qquad (54\cdot44)$$

Experimental measurements of the intensities of spectrum lines in emission and absorption give the empirical values of $\epsilon(n',n'')$ and $\alpha(n',n'')$, as the case may be.

54f. Spectroscopic Stability.—In using Eqs. (54·8), (54·9), and (54·10) for the calculation of the intensity of a spectrum line due to transitions between degenerate energy levels we can employ any set of orthogonal base functions which are eigenfunctions of the Hamiltonian of the molecule or atom under consideration. To pass from one such scheme to another, one must subject the matrices $^{(n',n'')}\mathbf{D}_x$, $^{(n',n'')}\mathbf{D}_y$, $^{(n',n'')}\mathbf{D}_z$ to canonical transformations which leave the diagonal sums involved in the above equations unchanged.

When a weak constant perturbing field is applied to an atom or molecule, the energy levels and associated spectrum lines are ordinarily broken up into fine structures. If the field is weak enough, it will be legitimate to compute the intensities of these lines using the zero-order wave functions of perturbation theory. These wave functions are always possible wave functions for the unperturbed atoms. The total intensity of all components of a multiple line created by the application of the field will be obtained in first approximation by a sum similar to that given in (54·8). Thus in the limit, as the field approaches zero, the sum of the intensities of the fine-structure components must approach the intensity of the unperturbed line. This is the principle of spectroscopic stability.

*54g. Magnetic Dipole and Electric Quadrupole Radiation.[1]—In computing the elements of $\mathbf{H}_1^{(t)}$ in Sec. 54d we treated the vector potential $\overrightarrow{a}$ as constant over the atomic system. This approximation gives

---

[1] For a correspondence principle treatment of electric quadrupole radiation, including a derivation of the selection and polarization rules in an external magnetic field, see A. Rubinowicz, *Zeits. f. Physik* **61**, 338 (1930); Rubinowicz and Blaton, *Ergeb. d. Exact. Naturwiss.* **11**, 181 (1932); Condon and Shortley, *T.A.S.*, Sec. 6⁴. Formula (7), p. 96 of *T.A.S.* is misleading as the dyadic $\mathfrak{N}$, *i.e.*, our Q, should be replaced by our $\tilde{Q}$ [*cf.* Eq. (54·65)].

the contribution of the electric dipole moment $\vec{D}$ to the transition probability. To carry the computation one step farther we imagine $\vec{\alpha}$ expanded in power series in x, y, z and keep the linear terms as well as the constant one, discarding the rest. The new terms give the contributions of the magnetic dipole and electric quadrupole radiation to the transition probability. They are practically always negligible in comparison with the dipole terms unless the latter vanish as a result of a selection rule. Hence we here neglect cross terms and calculate the transition probability due to the linear terms alone. In place of (54·16) we write

$$\lambda H_1 = -\frac{h}{2\pi i c}\sum_j \frac{e_j}{\mu_j}\left\{\left[x_j\frac{\partial\alpha_x}{\partial x} + y_j\frac{\partial\alpha_x}{\partial y} + z_j\frac{\partial\alpha_x}{\partial z}\right]\frac{\partial}{\partial x_j} + \left[x_j\frac{\partial\alpha_y}{\partial x}\right.\right.$$
$$\left.+ y_j\frac{\partial\alpha_y}{\partial y} + z_j\frac{\partial\alpha_y}{\partial z}\right]\frac{\partial}{\partial y_j} + \left[x_j\frac{\partial\alpha_z}{\partial x} + y_j\frac{\partial\alpha_z}{\partial y} + z_j\frac{\partial\alpha_z}{\partial z}\right]\frac{\partial}{\partial z_j}\Bigg\}. \quad (54·45)$$

with the understanding that the derivatives of $\vec{\alpha}$ are to be evaluated at the origin (*i.e.*, at the center of mass of the system).

H_1 is linear in the elements of the matrix

$$\mathbf{S} = \begin{Vmatrix} \dfrac{\partial\alpha_x}{\partial x} & \dfrac{\partial\alpha_x}{\partial y} & \dfrac{\partial\alpha_x}{\partial z} \\ \dfrac{\partial\alpha_y}{\partial x} & \dfrac{\partial\alpha_y}{\partial y} & \dfrac{\partial\alpha_y}{\partial z} \\ \dfrac{\partial\alpha_z}{\partial x} & \dfrac{\partial\alpha_z}{\partial y} & \dfrac{\partial\alpha_z}{\partial z} \end{Vmatrix}.$$

This matrix is at once resolvable into the sum of matrices $\mathbf{S}_s$ and $\mathbf{S}_a$ symmetric and antisymmetric respectively with respect to a transposition of rows and columns. With the aid of the formula, $\vec{\mathcal{H}} = \text{curl } \vec{\alpha}$, we obtain

$$\mathbf{S}_s = \begin{Vmatrix} \dfrac{\partial\alpha_x}{\partial x} & \dfrac{1}{2}\left(\dfrac{\partial\alpha_x}{\partial y} + \dfrac{\partial\alpha_y}{\partial x}\right) & \dfrac{1}{2}\left(\dfrac{\partial\alpha_x}{\partial z} + \dfrac{\partial\alpha_z}{\partial x}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial\alpha_x}{\partial y} + \dfrac{\partial\alpha_y}{\partial x}\right) & \dfrac{\partial\alpha_y}{\partial y} & \dfrac{1}{2}\left(\dfrac{\partial\alpha_y}{\partial z} + \dfrac{\partial\alpha_z}{\partial y}\right) \\ \dfrac{1}{2}\left(\dfrac{\partial\alpha_x}{\partial z} + \dfrac{\partial\alpha_z}{\partial x}\right) & \dfrac{1}{2}\left(\dfrac{\partial\alpha_y}{\partial z} + \dfrac{\partial\alpha_z}{\partial y}\right) & \dfrac{\partial\alpha_z}{\partial z} \end{Vmatrix};$$

$$\mathbf{S}_a = \begin{Vmatrix} 0 & -\dfrac{1}{2}\mathcal{H}_z & \dfrac{1}{2}\mathcal{H}_y \\ \dfrac{1}{2}\mathcal{H}_z & 0 & -\dfrac{1}{2}\mathcal{H}_x \\ -\dfrac{1}{2}\mathcal{H}_y & \dfrac{1}{2}\mathcal{H}_x & 0 \end{Vmatrix}. \quad (54·46)$$

In (54.45) we next replace each element of **S** by the sum of the corresponding elements of $\mathbf{S}_s$ and $\mathbf{S}_a$. H_1 is thereby resolved into the sum of two terms K_Q and $K_{\mathfrak{IC}}$ given by the equations

$$\lambda K_{\mathfrak{IC}} = -\frac{h}{4\pi ic}\sum_{j}\frac{e_j}{\mu_j}\left\{\mathfrak{IC}_x\left(y_i\frac{\partial}{\partial z_j} - z_i\frac{\partial}{\partial y_j}\right) + \mathfrak{IC}_y\left(z_i\frac{\partial}{\partial x_j} - x_i\frac{\partial}{\partial z_j}\right)\right.$$

$$\left. + \mathfrak{IC}_z\left(x_i\frac{\partial}{\partial y_j} - y_i\frac{\partial}{\partial x_j}\right)\right\}, \quad (54.47)$$

$$\lambda K_Q = -\frac{h}{2\pi ic}\sum_{j}\frac{e_j}{\mu_j}\left\{\frac{\partial\mathfrak{a}_x}{\partial x}x_i\frac{\partial}{\partial x_j} + \frac{\partial\mathfrak{a}_y}{\partial y}y_i\frac{\partial}{\partial y_j} + \frac{\partial\mathfrak{a}_z}{\partial z}z_i\frac{\partial}{\partial z_j}\right.$$

$$+ \frac{1}{2}\left(\frac{\partial\mathfrak{a}_x}{\partial y} + \frac{\partial\mathfrak{a}_y}{\partial x}\right)\left(x_i\frac{\partial}{\partial y_j} + y_i\frac{\partial}{\partial x_j}\right) + \frac{1}{2}\left(\frac{\partial\mathfrak{a}_y}{\partial z} + \frac{\partial\mathfrak{a}_z}{\partial y}\right)\left(y_i\frac{\partial}{\partial z_j} + z_i\frac{\partial}{\partial y_j}\right)$$

$$+ \frac{1}{2}\left(\frac{\partial\mathfrak{a}_z}{\partial x} + \frac{\partial\mathfrak{a}_x}{\partial z}\right)\left(z_i\frac{\partial}{\partial x_j} + x_i\frac{\partial}{\partial z_j}\right)\right\}. \quad (54 \cdot 48)$$

The coefficient of each component of $\overrightarrow{\mathfrak{IC}}$ is readily seen to be the negative of the corresponding component of the magnetic dipole moment vector operator $\overrightarrow{\mathfrak{M}}$ [*cf.* Eq. (49·13)]. Hence $K_{\mathfrak{IC}}$ reduces to the form

$$\lambda K_{\mathfrak{IC}} = -\overrightarrow{\mathfrak{IC}} \cdot \overrightarrow{\mathfrak{M}}. \quad (54 \cdot 49)$$

Here $\overrightarrow{\mathfrak{IC}}$ denotes the magnetic intensity at the origin and is a function of t only. The contribution of $K_{\mathfrak{IC}}$ to the matrix element $H_1^{(t)}(n,m)$ is accordingly

$$\frac{1}{\lambda}[\mathfrak{IC}_x(t)\mathfrak{M}_x(n,m) + \mathfrak{IC}_y(t)\mathfrak{M}_y(n,m) + \mathfrak{IC}_z(t)\mathfrak{M}_z(n,m)].$$

It follows from Eq. (53·15) that the corresponding contribution to $\mathbf{F}^{(1)}$, *i.e.*, $\mathbf{T}_1(\theta)^{(1)}$, is

$$F_{\mathfrak{IC}}{}^{(1)}(n,m) = \frac{2\pi i}{h}[\gamma_x(\nu_{nm})\mathfrak{M}_x(n,m) + \gamma_y(\nu_{nm})\mathfrak{M}_y(n,m) + \gamma_z(\nu_{nm})\mathfrak{M}_z(n,m)]$$

$$(54 \cdot 50)$$

where

$$\lambda\overrightarrow{\gamma}(\nu_{nm}) = \int_0^{\theta}\overrightarrow{\mathfrak{IC}}(t)e^{-2\pi i\nu_{nm}t}dt = \int_{-\infty}^{+\infty}\overrightarrow{\mathfrak{IC}}(t)e^{-2\pi i\nu_{nm}t}dt.$$

$\lambda\overrightarrow{\gamma}(\nu_{nm})$ is the Fourier transform (*cf.* p. 36 of Sec. 9a) of the field $\overrightarrow{\mathfrak{IC}}$ at the origin, evaluated for the frequency ν_{nm}. To derive the corresponding Einstein transition probability we have to take the square of the absolute value of $F^{(1)}(n,m)$ and average over all chaotic radiation fields consistent with a given macroscopic spectrum. Terms involving cross products such as $\gamma_x(\nu_{nm})\gamma_y(\nu_{nm})$ are eliminated by this averaging process. Moreover, since the mean-square intensity of the magnetic field for any

frequency is equal to the mean square of the electric field for the same frequency, we can replace $\vec{\gamma}(\nu_{nm})$ by $\vec{\phi}(\nu_{nm})$ in the final result. Thus Eq. (54·50) is equivalent to $F_{\mathfrak{X}}{}^{(1)}(n,m) = \frac{2\pi i}{h}\vec{\phi}(\nu_{nm}) \cdot \vec{\mathfrak{M}}(n,m)$. This is the same as (54·31) except for the substitution of the matrix of the magnetic dipole moment for the matrix of the electric dipole moment. We infer that the Einstein transition probability due to magnetic dipole radiation acting alone is

$$B_{n'\to n''} = \frac{8\pi^3}{3h^2 g_{n'}}\sum_{m'}\sum_{m''}|\vec{\mathfrak{M}}(n',m';\ n'',m'')|^2. \tag{54·51}$$

In applying this formula the magnetic moment of electron spin is to be added to the orbital moment here considered.[1]

Let us turn our attention next to the electric quadrupole moment term K_Q. In the following formulas all matrices are, like $\mathbf{H}_1^{(t)}$, of the Heisenberg type.

$$\begin{aligned}
\lambda\mathbf{K}_Q = {} & -\frac{1}{c}\Bigg[\frac{\partial\mathcal{Q}_x}{\partial x}\sum_j\frac{e_j}{\mu_j}\mathbf{x}_j\mathbf{p}_{xj} + \frac{\partial\mathcal{Q}_y}{\partial y}\sum_j\frac{e_j}{\mu_j}\mathbf{y}_j\mathbf{p}_{yj} + \frac{\partial\mathcal{Q}_z}{\partial z}\sum_j\frac{e_j}{\mu_j}\mathbf{z}_j\mathbf{p}_{zj} \\
& + \frac{1}{2}\Big(\frac{\partial\mathcal{Q}_x}{\partial y} + \frac{\partial\mathcal{Q}_y}{\partial x}\Big)\sum_j\frac{e_j}{\mu_j}(\mathbf{x}_j\mathbf{p}_{yj} + \mathbf{y}_j\mathbf{p}_{xj}) + \cdots\Bigg] \\
= {} & -\frac{1}{2c}\Bigg[\frac{\partial\mathcal{Q}_x}{\partial x}\frac{d}{dt}\sum_j e_j\mathbf{x}_j{}^2 + \frac{\partial\mathcal{Q}_y}{\partial y}\frac{d}{dt}\sum_j e_j\mathbf{y}_j{}^2 + \cdots \\
& + \Big(\frac{\partial\mathcal{Q}_x}{\partial y} + \frac{\partial\mathcal{Q}_y}{\partial x}\Big)\frac{d}{dt}\sum_j e_j\mathbf{x}_j\mathbf{y}_j + \cdots\Bigg]. \tag{54·52}
\end{aligned}$$

Let Q_{xx}, Q_{xy}, etc., denote the components of the electric quadrupole tensor $\sum_j e_j x_j{}^2$, $\sum_j e_j x_j y_j$, $\cdots$, Carrying out the differentiations with respect to t indicated in (54·52), we obtain

$$\begin{aligned}
\lambda K_Q(n,m) = {} & -\frac{\pi i}{c}\nu_{nm}\Bigg[\frac{\partial\mathcal{Q}_x}{\partial x}Q_{xx}^{(t)}(n,m) + \frac{\partial\mathcal{Q}_y}{\partial y}Q_{yy}^{(t)}(n,m) \\
& + \frac{\partial\mathcal{Q}_z}{\partial z}Q_{zz}^{(t)}(n,m) + \Big(\frac{\partial\mathcal{Q}_x}{\partial y} + \frac{\partial\mathcal{Q}_y}{\partial x}\Big)Q_{xy}^{(t)}(n,m) + \cdots\Bigg]. \tag{54·53}
\end{aligned}$$

and

$$F_Q^{(t)}(n,m) = -\frac{2\pi^2\nu_{nm}}{\lambda hc}\Bigg[\int_0^\theta\Big\{\frac{\partial\mathcal{Q}_x}{\partial x}Q_{xx}(n,m) + \cdots \\
\Big(\frac{\partial\mathcal{Q}_x}{\partial y} + \frac{\partial\mathcal{Q}_y}{\partial x}\Big)Q_{xy}(n,m) + \cdots\Big\}e^{2\pi i\nu_{nm}t}dt\Bigg]. \tag{54·54}$$

[1] *Cf.* H. C. BRINKMAN, Thesis, Leiden, 1932.

In this last equation the matrix elements of the quadrupole moment tensor are of the Schrödinger, time-free, type.

Let $G_{xx}(\nu)$, $G_{xy}(\nu)$, etc., denote the Fourier transforms

$$G_{xx}(\nu) = \int_{-\infty}^{+\infty} \left(\frac{\partial A_x}{\partial x}\right)_0 e^{2\pi i \nu t} dt; \quad G_{xy}(\nu) = \int_{-\infty}^{+\infty} \left(\frac{\partial A_x}{\partial y} + \frac{\partial A_y}{\partial x}\right)_0 e^{2\pi i \nu t} dt: \cdots$$

where the subscript zero indicates that the quantity in question is to be evaluated at the origin. Equations (53·17) and (54·54) now yield

$$\Phi_{n \to m} = \frac{4\pi^4 \nu_{nm}^2}{h^2 c^2} |G_{xx}(\nu_{nm})Q_{xx}(n,m) + G_{yy}(\nu_{nm})Q_{yy}(n,m) + \cdots$$
$$+ G_{xy}(\nu_{nm})Q_{xy}(n,m) + \cdots |^2. \quad (54\cdot55)$$

Taking mean values over a chaotic assemblage of isotropic radiation fields, we note that

$$\overline{|G_{xx}|^2} = \overline{|G_{yy}|^2} = \overline{|G_{zz}|^2}, \qquad \overline{|G_{xy}|^2} = \overline{|G_{yz}|^2} = \overline{|G_{zz}|^2},$$
$$\overline{G_{xx}G_{yy}{}^*} = \overline{G_{xx}{}^*G_{yy}} = \overline{G_{yy}G_{zz}{}^*} = \overline{G_{yy}{}^*G_{zz}} = \overline{G_{zz}G_{xx}{}^*} = \overline{G_{zz}{}^*G_{xx}}.$$

The mean values of such products as $\overline{G_{xx}G_{xy}{}^*}$, $\overline{G_{xy}G_{yz}{}^*}$ are zero. Since $\mathrm{div}\ \overrightarrow{a} = 0$, $|G_{xx} + G_{yy} + G_{zz}|^2$ vanishes. Hence $\overline{G_{xx}G_{yy}{}^*} = -\frac{1}{2}\overline{|G_{zz}|^2}$.

The writer is indebted to Prof. J. H. Van Vleck for the suggestion of the following scheme for relating $\overline{|G_{xy}|^2}$ to $\overline{|G_{zz}|^2}$. The latter quantity is invariant of a rotation of the axes. Let l, m, n denote the direction cosines of a new z axis, say z', with respect to the original axes x, y, z respectively. One can then equate $\overline{|G_{zz}|^2}$ with its transform $\overline{|G_{z'z'}|^2}$ and demand that the equation shall hold identically in l, m, and n. We obtain

$$\overline{|G_{zz}|^2} = \overline{|G_{zz}|^2}\{(l^4 + m^4 + n^4) - (l^2m^2 + m^2n^2 + n^2l^2)\}$$
$$+ \overline{|G_{xy}|^2}(l^2m^2 + m^2n^2 + n^2l^2).$$

Since $(l^2 + m^2 + n^2)^2 = 1$, the above relation requires that

$$\overline{|G_{xy}|^2} = 3\overline{|G_{zz}|^2}.$$

Equation (54·55) becomes

$$\Phi_{n \to m} = \frac{4\pi^4 \nu_{nm}^2}{h^2 c^2}\overline{|G_{zz}|^2}\Big\{|Q_{xx}|^2 + |Q_{yy}|^2 + |Q_{zz}|^2 + 3(|Q_{xy}|^2 + |Q_{yz}|^2$$
$$+ |Q_{zx}|^2) - \frac{1}{2}(Q_{xx}Q_{yy}{}^* + Q_{xx}{}^*Q_{yy} + Q_{yy}Q_{zz}{}^* + \cdots)\Big\}. \quad (54\cdot56)$$

Denoting the quantity $\sum_j e_j r_j^2$ by Λ, we have

$$|Q_{xx}(n,m) + Q_{yy}(n,m) + Q_{zz}(n,m)|^2 = |\Lambda(n,m)|^2.$$

With the aid of this relation we can reduce (54·56) to the form

$$\Phi_{n \to m} = \frac{6\pi^4 \nu_{nm}{}^2}{h^2 c^2} \overline{|G_{zz}(\nu_{nm})|^2} \Big\{ |Q_{xx}(n,m)|^2 + |Q_{yy}(n,m)|^2 + |Q_{zz}(n,m)|^2$$

$$+ 2|Q_{xy}(n,m)|^2 + 2|Q_{yz}(n,m)|^2 + 2|Q_{zx}(n,m)|^2 - \frac{1}{3}|\Lambda(n,m)|^2 \Big\}. \quad (54\cdot57)$$

As a final step in evaluating the Einstein transition probability for an isotropic radiation field we must relate $\overline{|G_{zz}(\nu)|^2}$ with the energy density $\rho(\nu)$. It is necessary for this purpose to describe the perturbing field as a superposition of plane waves. We accordingly express $\vec{\alpha}$ in the form of a multiple Fourier integral:

$$\vec{\alpha}(x,y,z,t) = \int\int\limits_{-\infty}^{+\infty}\int [\vec{A}_1(\vec{\sigma})e^{i\varphi} + \vec{A}_2(\vec{\sigma})e^{-i\varphi}]d\sigma_x d\sigma_y d\sigma_z \quad (54\cdot58)$$

where

$$\varphi = 2\pi[x\sigma_x + x\sigma_y + z\sigma_z - \nu t].$$

Then

$$\frac{\partial \alpha_z}{\partial z} = 2\pi i \int\int\int \sigma_z[A_{1z}(\vec{\sigma})e^{i\varphi} - A_{2z}(\vec{\sigma})e^{-i\varphi}]d\sigma_x d\sigma_y d\sigma_z.$$

Let $\vec{B}(\vec{\sigma})$ denote the vector $\vec{A}_1(\vec{\sigma}) - \vec{A}_2(\vec{\sigma})$. Then

$$\left(\frac{\partial \alpha_z}{\partial z}\right)_{t=0} = 2\pi i \int\int\int \sigma_z B_z(\vec{\sigma})e^{2\pi i \vec{r}\cdot\vec{\sigma}}d\sigma_x d\sigma_y d\sigma_z.$$

By Plancherel's theorem

$$\int\int\int\left(\frac{\partial \alpha_z}{\partial z}\right)^2_{t=0}dxdydz = 4\pi^2\int\int\int \sigma_z{}^2|B_z(\vec{\sigma})|^2 d\sigma_x d\sigma_y d\sigma_z$$

$$= \frac{4\pi^2}{c^3}\int_0^\infty \nu^2 d\nu \int\int \sigma_z{}^2|B_z(\vec{\sigma})|^2 d\Omega \quad (54\cdot59)$$

Here $d\Omega$ denotes an element of solid angle in σ-space.

We are interested solely in the field at the origin between $t = 0$ and $t = \theta$. According to Huygens' principle this field is determined by the field at $t = 0$ inside a sphere of radius $c\theta$ drawn about the origin as a center. The average value of $(\partial\alpha_z/\partial z)^2$ for this initial spatially distributed field can be identified with the average value of $(\partial\alpha_z/\partial z)^2$ at the origin for the time interval $0 < t < \theta$. We shall in fact identify the mean contribution of each portion of the spectrum to one of these average values with the corresponding mean contribution to the other. First of all

$$\frac{1}{\theta}\int_0^\theta\left(\frac{\partial \alpha_z}{\partial z}\right)^2_{\text{origin}}dt = \frac{2}{\theta}\int_0^\infty |G_{zz}(\nu)|^2 d\nu$$

$$= \frac{3}{4\pi c^3\theta^3}\int\int\int\left(\frac{\partial \alpha_z}{\partial z}\right)^2_{t=0}dxdydz. \quad (54\cdot60)$$

Combining (54·59) with (54·60), we obtain

$$\int_0^\infty |G_{zz}(\nu)|^2 d\nu = \frac{3\pi}{2c^6\theta^2} \int_0^\infty \nu^2 d\nu \int\int \sigma_z{}^2 |B_z(\vec{\sigma})|^2 d\Omega.$$

Equating the statistical mean values of the integrands for a chaotic assemblage of isotropic radiation fields, we have

$$\overline{|G_{zz}(\nu)|^2} = \frac{3\pi\nu^2}{2c^6\theta^2} \int\int \overline{\sigma_z{}^2 |B_z(\vec{\sigma})|^2} d\Omega. \tag{54·61}$$

A similar treatment of $\dfrac{1}{c}\dfrac{\partial \mathcal{Q}_z}{\partial t}$ yields

$$\frac{1}{\theta}\int_0^\theta \left(\frac{1}{c}\frac{\partial \mathcal{Q}_z}{\partial t}\right)^2_{\text{origin}} dt = \frac{2\lambda^2}{\theta}\int_0^\infty |\phi_z(\nu)|^2 d\nu = \frac{3}{4\pi c^3}\theta^3 \int\int\int \left(\frac{1}{c}\frac{\partial \mathcal{Q}_z}{\partial t}\right)^2_{t=0} dx\,dy\,dz,$$

$$\lambda^2\overline{|\phi_z(\nu)|^2} = \frac{2\pi\nu^2}{2c^6\theta^2}\int\int \overline{\sigma^2|B_z(\vec{\sigma})|^2} d\Omega. \tag{54·62}$$

Here $\phi_z(\nu)$ is defined by (54·30). In view of the work of Sec. 54d we can identify $\lambda^2\overline{|\phi_z(\nu)|^2}$ with $\dfrac{2\pi\theta}{3}\rho(\nu)$. Equations (54·61) and (54·62) now give

$$\overline{|G_{zz}(\nu)|^2} = \frac{2\pi\theta}{3}\rho(\nu)\frac{\int\int \overline{\sigma_z{}^2|B_z(\vec{\sigma})|^2} d\Omega}{\int\int \overline{\sigma^2|B_z(\vec{\sigma})|^2} d\Omega}. \tag{54·63}$$

Since div $\mathcal{Q}$ is zero, the vectors $\vec{A}_1$, $\vec{A}_2$, and $\vec{B}$ are perpendicular to $\vec{\sigma}$. If ψ denotes the angle between $\vec{B}$ and the plane through $\vec{\sigma}$ and the z axis, while χ denotes that between $\vec{\sigma}$ and the positive z axis, we have $B_z = |\vec{B}| \cos\psi \sin\chi$. Hence

$$\overline{|B_z|^2} = \overline{|\vec{B}|^2}\,\overline{\cos^2\psi}\,\sin^2\chi = \tfrac{1}{2}\overline{|\vec{B}|^2}\sin^2\chi.$$

$\overline{|\vec{B}|^2}$ will be independent of the orientation of the vector $\vec{\sigma}$. Thus

$$\overline{|G_{zz}(\nu)|^2} = \frac{2\pi\theta}{3}\rho(\nu)\frac{\int_0^\pi \sin^3\chi\,\cos^2\chi\,d\chi}{\int_0^\pi \sin^3\chi\,d\chi} = \frac{2\pi\theta}{15}\rho(\nu).$$

On removing unnecessary absolute value signs (54·57) now yields

$$\Phi_{n\to m} = \frac{4\pi^5\theta}{5h^2c^2}\rho(\nu_{nm})\nu_{nm}{}^2 \left\{ \left(Q_{xx}(n,m) - \frac{\Lambda(n,m)}{3}\right)^2 \right.$$

$$+ \left(Q_{yy}(n,m) - \frac{\Lambda(n,m)}{3}\right)^2 + \left(Q_{zz}(n,m) - \frac{\Lambda(n,m)}{3}\right)^2$$

$$\left. + 2Q_{xy}(n,m)^2 + 2Q_{yz}(n,m)^2 + 2Q_{zx}(n,m)^2 \right\}. \tag{54·64}$$

Let $\tilde{Q}(n,m)$ denote the tensor whose matrix is

$$\left\|\begin{matrix} Q_{xx} - \frac{1}{3}\Lambda & Q_{xy} & Q_{xz} \\ Q_{yx} & Q_{yy} - \frac{1}{3}\Lambda & Q_{yz} \\ Q_{zx} & Q_{zy} & Q_{zz} - \frac{1}{3}\Lambda \end{matrix}\right\|.$$

Let $|\tilde{Q}(n,m)|^2$ denote the sum of the squares of the nine elements of $\tilde{Q}(n,m)$. (54·64) reduces to

$$\Phi_{n \to m} = \frac{4\pi^5 \theta}{5h^2 c^2} \rho(\nu_{nm}) |\tilde{Q}(n,m)|^2.$$

Using the double index notation appropriate to degenerate energy levels and summing over all types of transition from one such level to another we derive the Einstein transition probability

$$B_{n' \to n''} = \frac{4\pi^5 \nu_{nm}{}^2}{5h^2 c^2 g_{n'}} \sum_{m'} \sum_{m''} |\tilde{Q}(n',m'; n'',m'')|^2. \tag{54·65}$$

The transition probabilities $B_{n'' \to n'}$ and $A_{n' \to n''}$ follow directly from the above with the aid of Eqs. (54·4) and (54·5).

55. SOME ELEMENTARY SELECTION RULES FOR ELECTRIC DIPOLE RADIATION

55a. The Harmonic Oscillator.—The usual selection rules which distinguish "allowed" and "forbidden" jumps and permit us to predict the nature of the radiation spectrum of an atomic system are derivable from the fundamental intensity formulas for dipole radiation, (54·9) and (54·10). Let us consider first the ideal linear oscillator of Sec. 20. We assume that the oscillator has a charge e and proceed to compute the electric moment matrix, which can be treated as a scalar if the motion is along one of the coordinate axes:

$$D(n,n') = \int_{-\infty}^{+\infty} \psi_n^* e x \psi_{n'} dx.$$

The substitution

$$\psi_n = c_n H_n(\xi) e^{-\frac{\xi^2}{2}}, \qquad \xi = x\sqrt{\frac{k}{h\nu_c}},$$

yields

$$D(n,n') = \frac{eh\nu_c}{k} c_{n'} c_n^* \int_{-\infty}^{+\infty} H_{n'}(\xi) \xi H_n(\xi) e^{-\xi^2} d\xi.$$

The recurrence formula (20·8) gives

$$D(n,n') = \frac{eh\nu_c}{2k} c_{n'} c_n^* \int_{-\infty}^{+\infty} H_{n'}(\xi)[H_{n+1}(\xi) + 2nH_{n-1}(\xi)]e^{-\xi^2} d\xi.$$

The normalization-orthogonality rule (20·10) shows at once that $D(n,n')$ vanishes unless $n' = n \pm 1$. Spontaneous jumps from one state to

another with emission of radiation will then occur only from any given level to the next lower in the series. The only frequency radiated is $(E_n - E_{n-1})/h = \nu_c$, the frequency of the classical oscillation. The intensity of emission per atom in the initial energy level E_n is

$$I_{n \to n-1} = \frac{64\pi^4}{3c^3} \nu_c{}^4 |D(n, n-1)|^2, \qquad (55\cdot1)$$

where $D(n, n-1)$ is readily reduced to

$$D(n, n-1) = e\left(\frac{nh\nu_c}{2k}\right)^{1/2} e^{i(\chi_{n-1} - \chi_n)} = e\sqrt{\frac{E_n - E_0}{2k}} e^{i(\chi_{n-1} - \chi_n)}. \quad (55\cdot2)$$

Here χ_n and χ_{n-1} are the phase constants of c_n and c_{n-1}. The reader will readily verify that the intensity of radiation so calculated reduces to the classical value if E_0 is neglected in comparison with E_n.

55b. Selection Rules for the Two-particle Problem.—In the case of a two-particle problem with a homogeneous external magnetic field we have learned from Sec. 49 that the wave functions in the field are to a first-order approximation, the same as those obtained by separating the variables in spherical coordinates with the z axis in the direction of the field. Hence the matrix elements of the electric moment calculated from these functions can be used for computing both the intensities and polarizations of the components of the Zeeman pattern and for the intensities of the unperturbed lines.

Denoting the eigenvalues of $\mathcal{L}_z$ and $\mathcal{L}^2$ as usual by

$$\left.\begin{aligned} \mathcal{L}_z' &= \frac{mh}{2\pi}, & m &= 0, \pm1, \cdots \pm l \\ (\mathcal{L}^2)' &= \frac{l(l+1)h^2}{4\pi^2}, & l &= 0, 1, 2, \cdots \end{aligned}\right\} \qquad (55\cdot3)$$

we shall prove the selection rules

$$m \to m, m \pm 1; \qquad l \to l \pm 1. \qquad (55\cdot4)$$

In appropriate spherical coordinates the classical formulas for the components of the electric moment are

$$D_x = er \sin\theta \cos\varphi, \qquad D_y = er \sin\theta \sin\varphi, \qquad D_z = er \cos\theta.$$

It is convenient to work with the linear combinations

$$\mathfrak{D} = D_x + iD_y = er \sin\theta\, e^{i\varphi},$$
$$\mathfrak{D}^\dagger = D_x - iD_y = er \sin\theta\, e^{-i\varphi},$$

instead of D_x and D_y themselves. $\mathfrak{D}^\dagger$ is then the adjoint of $\mathfrak{D}$. Evidently

$$|\vec{D}|^2 = |D_x|^2 + |D_y|^2 + |D_z|^2 = \tfrac{1}{2}[|\mathfrak{D}|^2 + |\mathfrak{D}^\dagger|^2] + |D_z|^2. \quad (55\cdot5)$$

Using the normalized wave functions of Sec. 28, we have

$$\begin{aligned}
D_z(n,l,m;\ n',l',m') &= \mathfrak{D}_r(n,l;\ n',l')Z_\theta(l,m;\ l',m')Z_\varphi(m;m'), \\
\mathfrak{D}(n,l,m;\ n',l',m') &= \mathfrak{D}_r(n,l;\ n',l')\mathfrak{D}_\theta(l,m;\ l',m')\mathfrak{D}_\varphi(m;m'), \\
\mathfrak{D}^\dagger(n,l,m;\ n',l',m') &= \mathfrak{D}_r(n,l;\ n',l')\ \mathfrak{D}_\theta{}^\dagger(l,m;\ l',m')\mathfrak{D}_\varphi{}^\dagger(m;m'),
\end{aligned} \right\} \quad (55\cdot6)$$

where

$$\mathfrak{D}_r(n,l;\ n',l') \equiv e\int_0^\infty R_{nl}R_{n'l'}r^3dr = e\int_0^\infty r\mathfrak{R}_{nl}\mathfrak{R}_{n'l'}dr, \qquad (55\cdot7)$$

$$\begin{aligned}
Z_\theta(l,m;l',m') &\equiv \int_0^\pi \Theta_{lm}\Theta_{l'm'}\sin\theta\cos\theta\,d\theta, \\
\mathfrak{D}_\theta(l,m;l',m') &\equiv \mathfrak{D}_\theta{}^\dagger(l,m;l',m') = \int_0^\pi \Theta_{lm}\Theta_{l'm'}\sin^2\theta\,d\theta,
\end{aligned} \right\} \quad (55\cdot8)$$

$$\begin{aligned}
Z_\varphi(m;m') &\equiv \frac{1}{2\pi}\int_0^{2\pi} e^{i(m-m')\varphi}d\varphi = \delta_{mm'}, \\
\mathfrak{D}_\varphi(m;m') &\equiv \frac{1}{2\pi}\int_0^{2\pi} e^{i(m-m'+1)\varphi}d\varphi = \delta_{m+1,m'}, \\
\mathfrak{D}_\varphi{}^\dagger(m;m') &\equiv \frac{1}{2\pi}\int_0^{2\pi} e^{i(m-m'-1)\varphi}d\varphi = \delta_{m-1,m'}.
\end{aligned} \right\} \quad (55\cdot9)$$

Equations (55·9) show that all components of the $\overrightarrow{\mathbf{D}}$ matrix vanish unless $m' = m$ or $m \pm 1$. Thus the m selection principle drops out of the basic formulas without computation.

The principle of selection for the azimuthal quantum number is derivable from an examination of the integrals Z_θ, $\mathfrak{D}_\theta$, $\mathfrak{D}_\theta{}^\dagger$. In evaluating each integral we give m' the value required to prevent the cofactor Z_φ, $\mathfrak{D}_\varphi$, $\mathfrak{D}_\varphi{}^\dagger$, as the case may be, in Eqs. (55·6) from vanishing. Thus, introducing the symbol $C_{l,m}$ for the normalizing factor

$$\left\{\frac{(2l+1)(l-|m|)!}{2(l+|m|)!}\right\}^{\frac{1}{2}}$$

of Θ_{lm}, we have

$$Z_\theta(l,m;\ l'm) = C_{l,m}C_{l',m}\int_{-1}^{+1} x\,P_{l,|m|}(x)P_{l',|m|}(x)dx. \qquad (55\cdot10)$$

Formula (F-10), Appendix F, permits us to express $xP_{l,|m|}(x)$ as a linear combination of $P_{l-1,|m|}$ and $P_{l+1,|m|}$. In view of the orthogonality equation (F-15) it follows that $Z_\theta(l,m;l',m)$ vanishes unless $l' = l \pm 1$. Similarly $\mathfrak{D}_\theta(l,m;l',m+1)$ and $\mathfrak{D}_\theta(l,m;l',m-1)$ vanish unless $l' = l \pm 1$. Thus all components of the electric-moment matrix vanish for transitions in which l changes by any amount other than ± 1. This proves the selection rule for the azimuthal quantum number l.

55c. Fine Structure and Polarization of Spectrum Lines in Simple Zeeman Effect.—Consider next the application of the theory of absorption to hydrogenic atoms without electron spin in a uniform magnetic

field of strength $\mathfrak{IC}$. Taking the z axis in the direction of the field we learn from Sec. 49c that the energy of the state

$$u_{n,l,m} = R_{n,l}(r)\Theta_{l,m}(\theta)e^{\frac{im\varphi}{\sqrt{2\pi}}}$$

is $E_n^{(0)} + mh\nu_{\mathfrak{IC}}$ where $\nu_{\mathfrak{IC}}$ is the Larmor frequency, $|e|\mathfrak{IC}/4\pi\mu c$, and $E_n^{(0)}$ is the unperturbed energy. Let $\nu_{n,n'}^{(0)}$ denote the frequency of the unperturbed emission and absorption line due to transitions between the states $u_{n,l,m}$ and $u_{n',l',m'}$. We assume that $E_n > E_{n'}$ so that

$$\nu_{n,n'}^{(0)} = \frac{(E_n - E_{n'})}{h}.$$

The frequency of the perturbed line due to this special transition is then $\nu_{n,n'}^{(0)} + (m - m')\nu_{\mathfrak{IC}}$. Every unperturbed line is accordingly split by the field into three components whose frequencies are $\nu^{(0)} - \nu_{\mathfrak{IC}}$, $\nu^{(0)}$, and $\nu^{(0)} + \nu_{\mathfrak{IC}}$, coming respectively from the three types of allowed jumps in which $m - m'$ has the values -1, 0, and $+1$.

These three components show three distinctive types of polarization, as in the classical theory of the simple Zeeman effect. In order to demonstrate this we rewrite Eq. (54·31) in the form

$$\lambda F^{(1)}\{k,k'\} = -\frac{2\pi i\lambda}{h}\left\{\frac{1}{2}[\phi_x(\nu_{kk'}) - i\phi_y(\nu_{kk'})]\mathfrak{D}(k,k') + \frac{1}{2}[\phi_x(\nu_{kk'})\right.$$
$$\left. + i\phi_y(\nu_{kk'})]\mathfrak{D}^\dagger(k,k') + \phi_z(\nu_{kk'})Z(k,k')\right\}. \quad (55\cdot11)$$

In the case of the central component of the absorption triplet $m' = m$ and the matrix components of $\mathfrak{D}$ and $\mathfrak{D}^\dagger$ vanish. Thus only the component of the electric force of the impressed wave in the z direction is effective in producing absorption of this frequency. If light is traveling in the direction of the lines of force, or, if it is traveling perpendicular to the lines of force and is polarized with the electric vector at right angles to the lines of force, there is no absorption. This is the characteristic of the central component in the classical theory.

Consider next the high-frequency component $\nu = \nu^{(0)} + \nu_{\mathfrak{IC}}$. This is associated with the matrix element $\mathfrak{D}^\dagger(n,l,m; n',l',m - 1)$, the corresponding matrix elements of $\mathfrak{D}$ and Z being zero. Equation (55·11) reduces to

$$F^{(1)}\{k,k'\} = -\frac{2\pi i}{h}\left\{\frac{1}{2}[\phi_x(\nu_{kk'}) + i\phi_y(\nu_{kk'})]\mathfrak{D}^\dagger(k,k')\right\}. \quad (55\cdot12)$$

If the light of frequency $\nu = \nu^{(0)} + \nu_{\mathfrak{IC}}$ is an approximate plane wave traveling perpendicular to the lines of force, say along the x axis, only the component polarized with the electric vector perpendicular to the field is effective in producing absorption. If the light of frequency

$\nu^{(0)} + \nu_{3C}$ travels along the lines of force and is circularly polarized in the sense of a left-hand screw the phase of the y component of the electric force will be 90° ahead of that of the x component and $\phi_y = +i\phi_x$. Hence $\phi_x + i\phi_y$ vanishes and the light is not absorbed. Maximum absorption for a given value of the intensity occurs when the light is circularly polarized in the sense of a right-hand screw so that

$$\frac{\phi_x + i\phi_y}{2} = \phi_x.$$

The low-frequency component has, of course, the reverse behavior with respect to the two types of circularly polarized light propagated in the direction of the lines of force. Thus the Zeeman pattern in absorption for our model atom has precisely the same properties as in the familiar classical theory of the simple Zeeman effect for the same model.

The polarization characteristics of the emission triplet can be derived from those of the absorption triplet with the aid of Kirchhoff's law or the principle of detailed balancing.

The selection rules for many-particle problems require a rather elaborate analysis and will be worked out in Chap. XIV after we have developed methods for dealing with electron spin.

CHAPTER XIII

INTRODUCTION TO THE PROBLEM OF ATOMIC STRUCTURE: ELECTRON SPIN

56. THE ATOMIC PROBLEM AS A TWO-PARTICLE PROBLEM

56a. The Empirical Basis for the Idealized Bohr Atom Model.—Our present quantum-mechanical theory of the extranuclear structure of the atom•is an outgrowth and refinement of the Bohr interpretation of the optical spectra of the atoms. A number of important questions which arise in both theories were settled by Bohr's early investigations. On this account our approach to the present theory is one which leans heavily on the historical development of the earlier work.

In the sketch of the Bohr theory in Sec. 46 it was pointed out that, whereas strictly speaking the theory is applicable only to systems whose classical motions are multiply periodic, the original Rutherford model of any of the many-electron atoms gives rise to classical motions which, in the vast majority of cases, are not multiply periodic, even when the radiation resistance forces are neglected. In fact it is easy to see that spontaneous ionization would almost certainly be the speedy fate of a classical Rutherford atom started off with the energy of an empirical stationary state. Hence it was necessary to introduce an idealization of the Rutherford model whose motions are multiply periodic.

In devising this substitute model Bohr was guided by the empirical data. The alkali metals formed a group of elements of primary interest. They were known to carry one easily detachable electron—the valence electron of chemistry. They have particularly simple spectra which—ignoring fine structure—suggest the motion of a two-particle system, or of a single electron in a central force field. The energy levels are grouped together in a number of series known to spectroscopists as sharp series, principal series, etc. The spectra show that transitions do not occur between the different levels of the same series, or between all different series. They are subject to a principle of selection, and if we assign the ordinal numbers l as follows,

Series	Sharp	Principal	Diffuse	Fundamental	. . . .
l	0 .	1	2	3	. . . . . . , ...

it turns out that the selection rule is exactly that derived in Sec. 55b for the azimuthal, or angular-momentum, quantum number in a two-

particle problem, *viz.*, $\Delta l = \pm 1$. The development of the theory will show that the ordinal number l can in fact be identified with the angular-momentum quantum number of the valence electron.

Each series of levels is composed of an apparently infinite number of members whose energies approach a well-defined limit as we ascend the series. To each level there is assigned an ordinal number n giving its place in the series and beginning with an initial positive integer n_0 which was formerly assigned somewhat arbitrarily, but which can be chosen in such fashion as to give n a theoretical significance. The symbols s, p, d, f, g, $\cdot \cdot \cdot$ have been assigned to the successive series arranged as above in the order of increasing l values. The individual states are indicated by ns, np, nd, $\cdot \cdot \cdot$. Thus $3p$ denotes the state corresponding to a principal-series energy level with the ordinal number 3.

Taking the limit of any one of these series as the zero level of energy—not otherwise defined by spectroscopic data—we can describe the series by one or the other of the following well-known empirical formulas:

Rydberg formula:

$$E_{nl} = -\frac{Nhc}{(n - a_l)^2}. \qquad n = n_0, n_0 + 1, \cdot \cdot \cdot \qquad (56\cdot1)$$

Ritz formula:

$$E_{nl} = -\frac{Nhc}{(n - a_l + b_l E_{nl})^2}. \qquad n = n_0, n_0 + 1, \cdot \cdot \cdot \qquad (56\cdot2)$$

Here N is the Rydberg constant, sensibly the same as for the hydrogen series, and a_l and b_l are empirical constants for any one series, but functions of l if we wish to think of all the series together. The position of the lowest energy level of any given series fixes the value of the difference $n_0 - a_l$ in the Rydberg formula for that series. The value of n_0 was originally chosen for each series so as to make the associated value of a_l as small numerically as possible. The Ritz formula is evidently the more accurate as it involves two adjustable constants and includes the Rydberg formula as a special case.

These energy-level formulas bear a striking resemblance to the energy formula for the hydrogen atom—a resemblance reinforced by the observation that for large values of l the quantities a_l and $b_l E_{nl}$ can be made very small in comparison with unity by proper choice of the corresponding minimum integer n_0. In fact there is a critical value of l for each alkali metal beyond which all energy levels are in close agreement with corresponding hydrogen levels. This rule is illustrated by Fig. 23, which shows a chart of the energy levels of the optical spectrum of sodium with the energy laid out as abscissa and the azimuthal quantum number as ordinate. Vertical dotted lines indicate the energy levels of hydrogen.

These facts suggested the possibility of developing an approximate account of the spectrum of such an atom as sodium on the hypothesis

that the different energy levels represent different motions of the valence electron, the other particles forming a more or less stable symmetrical core or atom body. The assumed symmetry and stability of the central group are empirically reasonable in view of the fact that this group contains the same number of electrons as the very stable neon atom which precedes sodium in the periodic table. A similar statement applies to every alkali atom. The simplest assumption was then that in effect the valence electron moves in a central force field compounded of the field of the nucleus and the average field of the core. If the motion of the nucleus is taken into consideration, one has a typical two-particle
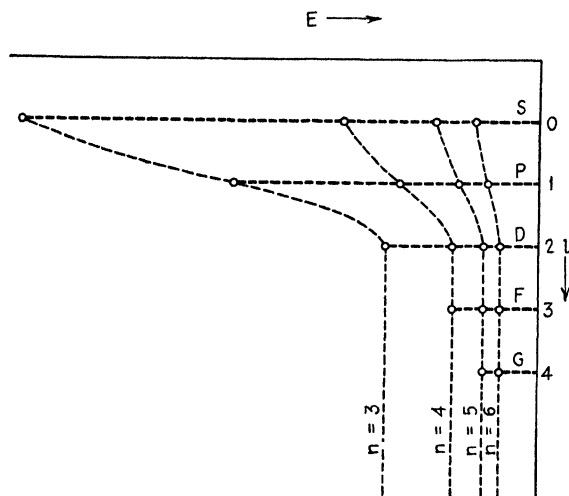
Fig. 23.—Diagram of the energy levels of the optical spectrum of sodium. (*After Bohr.*)

problem like that of Sec. 28. The substitution of this two-particle model for the actual complex atom implies a definite break with the third Bohr postulate, but such a break was *a priori* inevitable, as we have already seen. It was immediately justified by the light which it threw not only on the spectra of the alkalies but also on the general problem of atomic structure.

We shall not attempt here to follow in detail the application of Bohr's theory to this central-field problem but will develop the same idea by the method of wave mechanics. Our first objective will be the derivation of the Ritz formula (56·2) from our model.

The essential feature of this treatment of the sodium atom as a two-particle problem is that we neglect the details of the interaction between the inner electrons and the valence electron, considering only the average interaction. Translating this statement into the language of wave mechanics, we assume that to a first approximation the wave function of the atom can be resolved into the product of two factors, one of which

involves the coordinates of the valence electron only, while the other describes the state of the core and depends only on the coordinates of the inner group of electrons. We assume that the latter function is similar to the wave function of the normal state of the unperturbed ion obtained by removing the valence electron, and concern ourselves only with the former function.

The Schrödinger equation which determines the contribution of the valence electron to the atomic energy will then be of the form of the two-particle equation (28·1). In order to determine the qualitative character of the potential of the field in which the valence electron moves, we note that by elementary electrostatics the field of a uniform spherical shell of electricity is zero at interior points, and at exterior points is equal to the field of the same total charge when concentrated at the center of the shell. In calculating the potential we treat the core as a spherical distribution of charge of total amount $(Z - 1)e$, where Z is the atomic number. We shall suppose that practically all of this charge lies inside a radius r_1 of the order of magnitude of the kinetic-theory radius of the inert-gas atom of atomic number $Z - 1$. Then, if $V(r)$ is the net potential energy due to the nuclear charge and the core, we have

$$\frac{dV}{dr} = \frac{e^2}{r^2}, \qquad r > r_1 \tag{56·3}$$

and

$$\lim_{r \to 0} \left[r^2 \frac{dV}{dr} \right] = Ze^2. \tag{56·4}$$

There are a number of ways of estimating appropriate values of $V(r)$ for the interior of the core, of which the best known are the method of Thomas and Fermi[1] and the method of the Hartree self-consistent field.[2] For our present qualitative purpose, however, it will suffice to assume that $r^2 dV/dr$ varies linearly with r between the limits 0 and r_1. Figure 24 shows a comparison between the potential-energy curves for the hydrogen-atom and the sodium-atom model ($Z = 11$), the latter worked out on the above simple basis, assuming that, for sodium, r_1 is twice the radius a_0 of the innermost Bohr orbit for hydrogen.

The reason for this choice of r_1 is as follows. We have already observed that the energies of states of the alkali atoms with large values of the azimuthal quantum number l are very nearly coincident with

[1] L. H. THOMAS, *Proc. Cambridge Phil. Soc.* **23**, 542 (1926); E. FERMI, *Zeits. f. Physik* **48**, 73 (1928), **49**, 550 (1928). A full bibliography of the literature of the Thomas-Fermi and Hartree methods is given by L. Brillouin in the pamphlets "L'Atome de Thomas-Fermi" and "Les Champs 'Self-Consistents' de Hartree et de Fock," Paris, 1934.

[2] D. R. HARTREE, *Proc. Cambridge Phil. Soc.* **24**, 89, 111, 426 (1928); J. C. SLATER, *Phys. Rev.* **35**, 210 (1930); V. FOCK, *Zeits. f. Physik* **61**, 126 (1930).

corresponding hydrogenic energies. We shall refer to states of this kind as *hydrogenic states*, or *hydrogenic orbits*. We assume that in states of this kind the electron spends practically all its time in the Coulomb field region outside the core ($|\psi|^2$ has values inside the core very small compared with those outside). There will be an effective "distance of closest approach to the nucleus" for such states, which can be calculated as if the valence electron moved in a pure Coulomb field. The minimum distance of closest approach for any hydrogenic state gives an upper limit to the value of r_1 because any electron which penetrates the core is sure to have its energy affected by that fact.
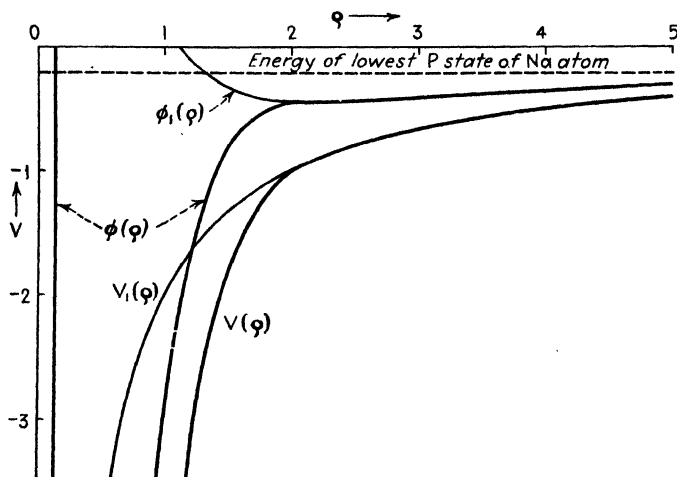
FIG. 24.—Approximate central-field potential energy for valence electron of sodium compared with potential energy for hydrogen atom. The abscissa ρ is r/a_0.

$V_1(\rho) = 2/\rho$: Potential energy of hydrogen atom in Hartree atomic units.

$V(\rho)$: Approximate potential energy for central-field model of sodium atom in atomic units.

$\phi_1(\rho) = V_1(\rho) + (1 + \frac{1}{2})^2/\rho$: Effective potential energy for hydrogen atom with $l = 1$.

$\phi(\rho) = V(\rho) + (1 + \frac{1}{2})^2/\rho$: Effective potential energy for sodium atom with $l = 1$.

Pauling and Goudsmit[1] have made a rough estimate of the distance of closest approach which gives the value a_0 (*cf.* p. 158) for principal-series orbits, $3a_0$ for diffuse-series orbits, etc. Empirically the energies of the principal-series states in sodium show marked deviations from the hydrogenic values, while the diffuse-series states show very slight deviations. Hence we can be reasonably sure that the radius of the core of the sodium atom is between the values a_0 and $3a_0$. The choice of $2a_0$ for r_1 used in preparing Fig. 24 is accordingly reasonable.

56b. Derivation of the Ritz Formula.—In order to place the empirical formulas (56·1) and (56·2) on a theoretical basis we separate the variables as in Sec. 28 and apply the B. W. K. method to the radial equation (28·19).

[1] L. PAULING and S. A. GOUDSMIT, *The Theory of Line Spectra*, p. 38, New York, 1930.

Following an ingenious scheme due to Bohr[1] we shall base our work on a comparison between the actual problem and the hydrogen-atom problem for the same energy and quantum number.

Let $\Phi(r,l)$ denote the effective potential energy $\left[V(r) + \dfrac{(l + \frac{1}{2})^2}{\kappa r^2} \right]$ for the valence electron of an alkali atom with azimuthal quantum number l [*Cf.* Secs. 21h and 28, Eq. (28·23)]. κ denotes the quantity $8\pi^2\mu/h^2$ as in previous chapters. Let $\Phi_{\rm H}(r,l)$ denote the corresponding effective potential energy $\left[-\dfrac{e^2}{r} + \dfrac{(l + \frac{1}{2})^2}{\kappa r^2} \right]$ for a hydrogen atom. The expressions for the classical local radial momentum in the two cases are then

$$p(r,l,E) = \sqrt{2\mu(E - \Phi)}, \qquad p_{\rm H}(r,l,E) = \sqrt{2\mu(E - \Phi_{\rm H})}.$$

We neglect the slight difference in the mass coefficients μ appropriate to the two different problems.

The condition for the location of the discrete energy levels of the problem in hand is

$$J = \oint p(r,l,E)dr = (v + \frac{1}{2})h. \qquad v = 0, 1, 2, \cdots \qquad (56·5)$$

The corresponding condition for the hydrogenic problem is

$$J_{\rm H} = \oint p_{\rm H}(r,l,E)dr = (v_{\rm H} + \frac{1}{2})h. \qquad v_{\rm H} = 0, 1, 2, \cdots \qquad (56·6)$$

The latter condition is known to be equivalent to the familiar equation (*cf.* Sec. 29a),

$$E = -\frac{Nhc}{(v_{\rm H} + l + 1)^2}, \qquad (56·7)$$

whether $v_{\rm H}$ is an integer or not. Let us now define a continuous function $n^*(E)$ (here the asterisk does *not* denote the complex conjugate) by means of the equation

$$E = -\frac{Nhc}{n^{*2}}. \qquad (56·8)$$

For all values of E, whether quantized or not, Eqs. (56·6), (56·7), and (56·8) lead to the relation

$$J_{\rm H}(E,l) = \oint p_{\rm H}(r,l,E)dr = (n^* - \frac{1}{2} - l)h. \qquad (56·9)$$

Replacing v in (56·5) by its value in terms of l and the total quantum

[1] N. BOHR, *Proc. London Phys. Soc.* **35**, 296 (1923). *Cf.* also J. H. Van Vleck, *Quantum Principles and Line Spectra*, pp. 110–114, Washington, 1926. The above discussions are based on the Bohr theory. For a pure wave-mechanics treatment see J. C. Slater, *Phys. Rev.* **31**, 333 (1928).

number n ($i.e.$, $v + l + 1$), we obtain for the quantum condition fixing the energy levels of the alkali-atom model

$$J(E,l) = \oint p(r,l,E)dr = (n - \tfrac{1}{2} - l)h. \qquad n = l + 1, l + 2, \cdots \tag{56·10}$$

The next step in the argument is to form the difference ΔJ between the phase integrals J and J_{H} for the alkali model and the hydrogen atom, using the same energy and azimuthal quantum number for each. In this way we obtain

$$\Delta J(E,l) = \oint p(r,l,E)dr - \oint p_{\mathrm{H}}(r,l,E)dr = (n - n^*)h. \tag{56·11}$$

E becomes an approximate eigenvalue of the Hamiltonian of the alkali model if n is chosen to be any positive integer, and l is given a positive integral value or set equal to zero.

Whereas J and J_{H} become infinite as E approaches zero, ΔJ is a slowly varying function of the energy throughout the range of energies encountered in a spectrum series (this observation is the key to the method). The contributions to the two integrals of the portions of the "orbits" outside the alkali core may be very large if the energy is small numerically, but they are equal in magnitude and cancel each other. The contributions of the portions of the two orbits inside the core ($r < r_1$) are nearly independent of the energy. This is especially true of the contribution to J, which is much greater than the corresponding contribution to J_{H}. These statements will be evident on examination of the effective potential-energy curves for principal-series energy levels shown in Fig. 24. In this figure a horizontal line indicates the lowest energy level of the principal series of sodium, and gives a basis for the graphical evaluation of $p(r,l,E)$ and $p_{\mathrm{H}}(r,l,E)$.

We conclude that $\Delta J/h$, $i.e.$, $n - n^*$, should be roughly constant and accurately representable as a linear function of E in the range of energies included in the actual series in question. Treating $n - n^*$ as constant, we get the Rydberg formula; writing

$$n - n^* = \frac{\Delta J}{h} = a_l - b_l E, \tag{56·12}$$

and substituting the resulting expression for n^* into (56·8) we obtain a formula of the Ritz type (56·2) with values of a_l and b_l which have theoretical significance. Using the theoretical value of $n - n^*$, we can choose n_0 so that the parameter n of Eqs. (56·1) and (56·2) becomes identical with the total quantum number $v + l + 1$.

The quantity $n - n^*$ is called the *quantum defect* since it defines the difference between the actual energy and the ideal energy of the corresponding state of the hydrogen atom. It is readily evaluated in first approximation from any assumed potential-energy curve when the energy

of the state in question is known experimentally. A graphical evaluation of ΔJ will suffice for this purpose. From the approximate value of $n - n^*$ obtained in this way and the experimental value of n^* we obtain a trial value of n which should be approximately integral, if our potential-energy curve is not too bad. We then identify the true total quantum number of the state in question with the integer which lies closest to the trial value. Thus in the case of the lowest energy level of the principal series of sodium a graphical integration yields the value 1.08 for the quantum defect. The empirical n^* is 2.12 and the trial value of n is 3.20. We conclude that the "true" value of the total quantum number n for this energy level is 3. The rather large deviation of the trial value of n from the nearest integer indicates that we have chosen too large a value for r_1.

The quantity b_l is readily worked out from the equation

$$b_l = -\frac{1}{h}\frac{\partial(\Delta J)}{\partial E} = \frac{\mu}{h}\left[\oint\frac{dr}{p_{\mathrm{H}}} - \oint\frac{dr}{p}\right]. \tag{56·13}$$

Classically μ/p is equal to the reciprocal of the velocity. Hence $\oint dr/p$ is a periodic time and $b_l h$ can be interpreted as the difference in the classical radial-oscillation periods for the laws of force of the sodium model and the hydrogen atom.

57. THE BOHR ASSIGNMENT OF ELECTRONIC QUANTUM NUMBERS

57a. The Quantum Numbers of the Valence Electrons in the Spectra of the Alkalies and Alkaline Earths.—In the absence of precise knowledge regarding the potential function $V(r)$ we cannot expect an accurate theoretical computation of the quantum defect. Even a rough estimate is useful, however, as a guide in making a correct assignment of the true total quantum number to each energy level. This is the more important inasmuch as the minimum value of n for each series is *not* equal to the theoretical value $l + 1$. In the case worked out in Sec. 56*b*, for example, n_0 turns out to be 3 when $l + 1 = 2$.

Some information regarding the n_0 values for different series can be obtained from a direct examination of energy-level diagrams like that shown in Fig. 23, thus avoiding the integrations of (56·11). We can assume that the total quantum number of each of the non-penetrating hydrogenic orbits is the same as that of the corresponding energy level of hydrogen. Furthermore, it follows from (56·11) that a_l *must increase as l decreases. Hence the energy must increase and decrease with l as we move along a line connecting states of the same total quantum number in a diagram in which energy and azimuthal quantum number are laid out as orthogonal coordinates.* In Fig. 23 a line of constant total quantum number must have a negative slope. This fact and the assignment of total quantum numbers for the hydrogenic orbits ($l \geqslant 2$) suffice for the

determination of minimum values of n_0 for each series. Thus in the sodium case, n_0 must be *at least* 3 for both the principal and the sharp series. A similar examination of the term system of lithium shows that the total quantum number for the normal state of this element (the lowest sharp-series state) must be *at least* 2. More generally, if we start with hydrogen and go up the series of elements with a single valence electron, we find that the total quantum number of the valence electron in the normal state increases by at least unity as we pass from each element to its successor. A systematic theoretical evaluation of the quantum defects, such as that carried through by Van Urk,[1] shows clearly that the actual increase in the total quantum number of the valence electron in its normal minimum energy state, as we pass from one alkali to the next in the series, must be exactly unity, and not a larger integer.

When this increase in the minimum total quantum number with the atomic number was first discovered there was no theoretical principle to account for it. States of lower total quantum number which were expected theoretically in the heavier atoms were not found experimentally. Their absence was a brute empirical fact which gave rise to the first elementary form of the Pauli exclusion principle.

Valuable supplementary evidence of an exclusion principle which prevents the occurrence of expected states of low energy is obtained from an examination of the spectra of the elements of the next two columns of the periodic table. Consider first the group of elements with two loose valence electrons consisting of helium and the alkaline-earth metals, beryllium, magnesium, calcium, etc. (*cf.* Thomsen diagram of periodic table on p. 489). The strong lines in the spectra of these elements—to be specific we shall focus our attention on the case of magnesium—come primarily from two systems of energy levels each similar in general character to the sodium system but differing in fine structure. For the present we ignore this fine structure except insofar as we apply the conventional designations "singlet" and "triplet" to the two systems under consideration. The same principles of selection apply to transitions inside any one of these systems as in the case of sodium. (Transitions from one system to the other are weak.) Consequently the levels can be classified like those of sodium. We are at once driven to interpret each system as due to the multiplicity of "excited" states of one of the valence electrons moving in a central force field due to the nucleus and a core composed of the remaining electrons.

It was at first difficult to understand how the core could have spherical symmetry in this case. One would suppose that the inner core of a magnesium atom obtained by removing *both* valence electrons would

---

[1] A. Th. Van Urk, *Zeits. f. Physik* **13**, 268 (1923). *Cf.* also M. Born. *loc. cit.*, pp. 195–198; Pauling and Goudsmit, *loc. cit.*, p. 40,

have the same spherical symmetry as the core of the sodium atom—each having the same number of electrons as a neon atom. Adding one valence electron in its lowest energy state should give a complete core for a magnesium atom with the same symmetry as a complete sodium atom. These conjectures are in fact verified by an examination of the spectrum of ionized magnesium. But according to the Bohr theory the normal state of the valence electron of sodium had one unit of angular momentum and could not give rise to a symmetrical average force field. Happily, this difficulty disappears when we adopt the point of view of wave mechanics which assigns zero angular momentum and a spherically symmetric wave function to the valence electron of sodium in its normal state.

The existence of the two different systems of energy levels suggests two alternative central force fields—a possibility we are not yet ready to interpret. This puzzling feature of the situation does not prevent us, however, from reading the primary lesson of the magnesium energy-level diagram, which is that the normal state to which one of the valence electrons returns after a period of excitation has the quantum numbers $n = 3$, $l = 0$. Using the conventional spectroscopic notation, we designate this state by the symbol 3s where the figure 3 specifies the value of n and the letter s indicates that $l = 0$. As the normal state of the inner valence electron is proved by the spark spectrum to be of the 3s type also, we see that it must be possible to have two electrons in s states of the same total quantum number.

Objection may be taken to the above statement on the ground that when two electrons are added in succession to a doubly charged ion, the fact that the first goes into a 3s state when it is the only valence electron does not prove that it remains in a 3s state when a second valence electron is attached with comparable firmness. The question arises whether as the second electron passes from one state to another of lower energy (stronger binding) it does not perturb the inner electron more and more violently until in the end the 3s nomenclature with its implication of a plane orbit in a central force field is no longer applicable. In fact from the standpoint of the Bohr theory two coexisting 3s elliptic orbits in the same atom are very hard to imagine and would certainly violate the third Bohr postulate.

Nevertheless, it does make sense to say that both the valence electrons of magnesium in its normal state occupy equivalent 3s states. Speaking more generally, it is possible to validate the practice of assigning to each of the inner electrons of any atom, as well as to the valence electrons, a pair of quantum numbers n, l appropriate to a central force field. In order to justify these statements we shall turn aside from the main line of argument to consider this question of the assignment of quantum numbers to the individual electrons from the standpoint of wave mechanics.

57b. Perturbation Theory and the Significance of an Assignment of Quantum Numbers to Inner Electrons.—In the chapter on perturbation theory we have seen that two different eigenvalue-eigenfunction problems A and B can frequently be related by means of a continuous series of interpolation problems whose Hamiltonian $\tilde{H}(\lambda)$ degenerates into that of A for one value of λ and into that of B for another value. In this case the interpolation eigenvalues $\tilde{E}_n(\lambda)$ and eigenfunctions $\tilde{\psi}_n(x,\lambda)$ form a continuous bridge between corresponding eigenvalues and eigenfunctions of A and B. If the problem A has a higher symmetry and a higher degree of degeneracy than B, the energy levels will split as λ passes from the value appropriate to A, say zero, toward the value appropriate to B, say unity. Hence each energy level of A will be correlated with a multiplicity of energy levels of B. On the other hand, unless B has also a special type of symmetry which is lacking in A and in the interpolation problem, there will be no corresponding splitting of energy levels as we pass from $\lambda = 1$ toward $\lambda = 0$. Thus, except for accidental degeneracy, which we can ignore, we can say that each energy level of the less symmetric problem B is correlated with a single definite energy level of A. If the problem B has degenerate eigenvalues, though less degenerate than those of A, it will be because there are two or more independent observables which commute with both H_B and $\tilde{H}$. $\tilde{H}$ must unite with one or more observables $\alpha_1, \alpha_2, \cdots$ to form a complete set of independent commuting observables. Each simultaneous eigenfunction of $\tilde{H}$ and the α's is nondegenerate and forms a perfectly definite connection between corresponding eigenfunctions of the problems A and B.

It will now be proved that if we identify the Hamiltonian of B with that of our atomic model, we can identify A with a parallel problem in which each electron moves in a central force field independent of the others. The potential energy for a neutral atom with nuclear charge Ze is

$$V_B = -\sum_j \frac{Ze^2}{r_j} + \sum_{j,k} \frac{e^2}{r_{jk}}, \tag{57.1}$$

where r_j is the distance of the jth electron from the nucleus, r_{jk} is the distance between electrons j and k, and the second sum is to be extended over all pairs of values of j and k [*cf.* Eq. (32.1)]. Let the potential energy of the problem A be

$$V_A = \sum_j V_0(r_j), \tag{57.2}$$

and let that of the interpolation problem be

$$\tilde{V} = V_A + \lambda(V_B - V_A). \tag{57.3}$$

In the limiting problem A there is no interaction between the electrons and each moves independently of the others. Eigenfunctions of this problem are obtained by separation of the variables as follows. Let

$$\psi^{(A)} = \prod_j \psi_j(\vec{r_j}), \qquad E^{(A)} = \sum_j E_j^{(0)}, \tag{57.4}$$

where each ψ_j is an eigenfunction of

$$\nabla^2 \psi_j + \kappa(E - V_0(r_j))\psi_j = 0 \tag{57.5}$$

with the eigenvalue $E_j^{(0)}$. Then $\psi^{(A)}$ is an eigenfunction of the central-field Hamiltonian H_A. Thus $\bar{V}$ yields the desired connection between the problem of the atom under consideration and a related central-field problem of the type encountered in Secs. 28 and 29.

It is, of course, legitimate to assign a definite set of quantum numbers to each electron in a suitably chosen eigenfunction of problem A. For example we may identify each of the ψ_j's with one of the solutions of (57.5) obtained by separating the variables in spherical coordinates. Thus we choose for $\psi^{(0)}$ the expression

$$\psi_{n_1, l_1, m_1, n_2, \cdots}^{(A)} = \prod_j \psi_{n_j, l_j, m_j}(\vec{r_j}), \tag{57.6}$$

$$\psi_{n_j, l_j, m_j} = (2\pi)^{-\frac{1}{2}} R_{n_j, l_j}(r_j)\Theta_{l_j, m_j}(\theta_j)e^{im_j\varphi_j}. \tag{57.7}$$

It will be convenient to simplify the discussion by eliminating the continuous spectrum of both problems A and B through the introduction of the boundary condition that ψ shall vanish when any electron reaches a sphere of very large radius ρ drawn about the nucleus as a center. The modification produced in the experimentally available discrete eigenvalues and their eigenfunctions will be negligible. Under these circumstances a complete set of eigenfunctions of problem A is obtained by forming all possible products of the above type. The energy of each individual solution is

$$E^{(A)} = \sum_j E_{n_j, l_j}^{(0)}. \tag{57.8}$$

Accordingly we have two kinds of degeneracy in problem A, *viz.*, degeneracy due to the fact that the quantum number m_j associated with n_j and l_j can be chosen in $2l_j + 1$ different ways without affecting the energy, and permutation degeneracy due to the fact that an exchange of the quantum numbers associated with any pair of electrons does not affect the resultant energy. Such an exchange of quantum numbers is, of course, equivalent to a permutation of the coordinates of the different electrons (*cf.* Sec. 40*d*).

Owing to the degeneracy of the eigenfunctions of problem A, it is not true that *every* eigenfunction is factorable into a product of one-

electron central-field wave functions. On the contrary there is an infinite variety of linear combinations of product functions possible and such linear combinations are not product functions. Hence it is *not* literally correct to say that any eigenfunction of A assigns to each electron a private set of central-field quantum numbers. In fact one can readily prove that just those eigenfunctions of A which form the zero-order approximate eigenfunctions of B are *not* product functions and do not give each electron a definite set of individual quantum numbers.

On the other hand, it follows from (57·8) that every eigenfunction of A is a linear combination of product functions in which the same set of n,l values appears in each term. Let us take, for example, the very simple case in which there are just two electrons, as in the neutral helium atom. Consider the special problem A energy level

$$E^{(A)} = E_{1,0}^{(0)} + E_{2,1}^{(0)}.$$

A typical product eigenfunction would be

$$\psi_a^{(A)} = \psi_{1,0,0}(\vec{r_1})\psi_{2,1,0}(\vec{r_2}).$$

The degeneracy with respect to m_j gives two other independent eigenfunctions, *viz.*,

$$\psi_b^{(A)} = \psi_{1,0,0}(\vec{r_1})\psi_{2,1,1}(\vec{r_2}), \qquad \psi_c^{(A)} = \psi_{1,0,0}(\vec{r_1})\psi_{2,1,-1}(\vec{r_2}).$$

Permutation degeneracy adds three more wave functions obtained by interchanging the symbols $\vec{r_1}$ and $\vec{r_2}$ in the above expressions. In any linear combination of these six product functions we have the same pair of pairs of n,l values, *viz.*, 1,0 and 2,1. Thus every eigenfunction of $E^{(A)}$ and every corresponding zero-order eigenfunction of problem B can be regarded as a description of a mixture of states in each of which one electron occupies a $1s$ state, while the other occupies a $2p$ state. Hence it is reasonable to speak of this energy level as one having one $1s$ electron and one $2p$ electron, even though one cannot say which is which. In the general case of an atom with any number of electrons one can say that the eigenfunctions of problem A and the zero-order eigenfunctions of problem B are linear combinations of product functions in each of which the same number of electrons is assigned to every pair of values of n and l.

In the higher order approximations to problem B there is a mixing of eigenfunctions of A for different unperturbed energy levels. Nevertheless every energy level of B is correlated with a unique corresponding level of A. Hence it is reasonable to ascribe to the levels of B, designations which specify the A levels to which they belong. Thus we can speak of an atomic-energy level (B level) as characterized by a definite set of pairs of individual electron quantum numbers n,l when what we

mean is that the corresponding level of the modified problem A has eigenfunctions involving this set of quantum numbers and no others. (It is easy to see that no modification of problem A due to changing $V_0(r)$ can alter the n,l values of the A level correlated with any given B level.)   This nomenclature will be particularly appropriate if it is possible to choose the potential function $V_0(r)$ so that problem A really bears a close resemblance to problem B.

Actually it is possible in various ways to pick out forms for $V_0(r)$ which make the solution of problem A a useful, though crude, first approximation to the solution of B.   In fact theoretical studies of atoms and atomic spectra are largely made by perturbation methods using a problem of the A type as the starting point.   Such an initial approximation ignores the forces on any one electron due to the specific instantaneous positions of its companions, and substitutes a central force field which can be identified roughly with the average field due to the remaining electrons acting on the one under consideration when the latter is located at the given distance from the nucleus.   The problem of determining the best form for $V_0(r)$ is similar to the problem of determining $V(r)$ in Sec. 56a.   One can begin with an estimate of the average charge distribution in successive spherical shells drawn about the nucleus and obtained from the Fermi-Thomas statistical theory,[1] or from X-ray scattering experiments.   If this charge is thought of as uniformly distributed over each shell, the corresponding electrostatic potential function $\varphi(r)$ is readily computed. $\varphi(r)$ is the potential of Z electrons in the case of a neutral atom, whereas what we want is the potential of the $Z - 1$ electrons with which the electron under consideration is interacting.   Hence $\varphi(r)$ can be multiplied by $(Z - 1)/Z$, and $V_0(r)$ identified with

$$V_0(r) = -\frac{Ze^2}{r} + \frac{(Z - 1)e\varphi(r)}{Z}.$$

As there is an infinite variety of central-field problems correlated with the same fundamental B problem, it is possible to make different choices of $V_0(r)$ for different purposes.   Suppose, for example, that we desire to make computations of the relative energies and transition probabilities of the excited states of a sodium atom responsible for the optical series spectrum.   In this case a natural and simple procedure is to choose suitable, fixed, low quantum numbers for the 10 inner electrons which belong to the "core" of the atom and give the eleventh electron various pairs of quantum numbers representing states of energy equal to, or higher than, $E_{3,0}^{(0)}$.   The relative energies of the different states in question will then differ as a result of the different energies of the outer valence electron only, the contributions of the inner electrons being the

[1] *Cf.* footnote 1, p. 477.

same in all cases. The wave functions specified do not take into account permutation degeneracy and the Pauli principle. Nevertheless, in the interests of simplicity we here assume our right to compute transition probabilities from them. The assumption will be justified in Sec. 64b. The components of the electric-moment matrix worked out in this way depend only on the wave functions of the valence electron and are independent of the factors in the total wave function representing the inner electrons. To prove this statement we note that each matrix element of, say, the X component of the electric moment, is the sum of terms each involving the x coordinate of one electron. In a transition of the type under consideration in which only one electron jumps, the orthogonality of the two wave functions of the jumping electron eliminates all terms except the one involving the x coordinate of that particular electron. This term reduces at once to the matrix element computed from the wave functions of the valence electron alone. Thus for the special purpose under consideration it is unnecessary to introduce either the energies or the wave functions of the core electrons and we can choose $V_0(r)$ solely with the purpose of getting a correct representation of the behavior of the valence electron.

The general type A problem for the sodium atom reduces in the above case to the single-electron problem of Sec. 56b. The latter can also be regarded as the result of attempting to secure an approximate solution of the primary type B problem of the sodium atom in the form of a product of a wave function for the valence electron and a wave function for the 10 core electrons. Its success is to be attributed largely to the lack of serious overlapping of these two functions.

We conclude that the quantum numbers deduced for the normal state of the valence electron from the empirical series spectrum and the theory of Sec. 56 are, in fact, the same as the quantum numbers to be assigned to this electron on the basis of the general A problem scheme just developed. Let us now suppose that an atom of atomic number Z is transformed into an atom of atomic number $Z + 1$ by first increasing its nuclear charge gradually by $|e|$ and then adding an additional electron. The question arises whether, or not, the quantum numbers of the inner group of Z electrons in the new atom are necessarily the same as those of the group of Z electrons in the initial atom. In order to give a partial answer we note that the increase in the nuclear charge, which constitutes the first stage of the transformation, must inevitably transform each energy level of the initial atom into a corresponding level of the positive ion of the final atom having the same quantum numbers. The normal state of any atom or ion being the state of lowest energy, it is clear that the normal state of the atom Z will carry over into the normal state of the ion $(Z + 1)^+$ if the energetic order of the lowest levels is not altered by the increase in nuclear charge. Such a change in the energetic

order of the lowest levels should be a rare phenomenon, although it is known to occur at certain points in the periodic table. It is made evident, when it does occur, by a comparison of the spectra of the atom Z and the ion $(Z + 1)^+$. A shift in the quantum numbers of one or more of the inner electrons accompanying the binding of an additional outer electron is also not to be ruled out on purely *a priori* grounds. We can say, however, that such a shift should be unusual and must involve a

```
                                        ┌ 55 Cs ─────── 87?
                                        ┌ 56 Ba ─────── 88 Ra
                                        │ ┌57 La┐       ┌89 Ac┐
                                        │ │58 Ce│       │90 Th│
                                        │ │59 Pr│       │91 Pa│
                                        │ │60 Nd│       │92 U │
                                        │ │61 ? │       └─────┘
                      19 K ──── 37 Rb   │ │62 Sm│
                      20 Ca ─── 38 Sr   │ │63 Eu│
                      ┌21 Sc┐ ─┌39 Y ┐  │ │64 Gd│
                      │22 Ti│ ─│40 Zr│  │ │65 Tb│
                      │23 V │ ─│41 Nb│  │ │66 Dy│
          ┌3 Li ──11 Na│24 Cr│ ─│42 Mo│  │ │67 Ho│
          │4 Be ──12 Mg│25 Mn│ ─│43 Ma│  │ │68 Er│
          │5 B  ──13 Al│26 Fe│ ─│44 Ru│  │ │69 Tu│
    1 H ┤ 6 C  ──14 Si│27 Co│ ─│45 Rh│  │ │70 Yb│
    2 He│ 7 N  ──15 P │28 Ni│ ─│46 Pd│  │ │71 Lu│
          │8 O  ──16 S │29 Cu─── 47 Ag    │72 Hf│
          │9 F  ──17 Cl│30 Zn─── 48 Cd    │73 Ta│
          └10 Ne ──18 Ar│31 Ga─── 49 In    │74 W │
                      │32 Ge─── 50 Sn    │75 Re│
                      │33 As─── 51 Sb    │76 Os│
                      │34 Se─── 52 Te    │77 Ir│
                      │35 Br─── 53 I     │78 Pt│
                      └36 Kr─── 54 X     └ 79 Au
                                          ┌ 80 Hg
                                          │ 81 Tl
                                          │ 82 Pb
                                          │ 83 Bi
                                          │ 84 Po
                                          │ 85 ?
                                          └ 86 Rn
```

FIG. 25.—The Thomsen-Bohr diagram of the periodic table. Short-life radioactive elements of atomic number greater than 92 have been discovered recently by Fermi and collaborators.

multiple electron jump in the binding process which would be readily recognizable from an analysis of the spectrum produced by that process. Actually no such shifts are known experimentally. It is accordingly legitimate in general to use the optical spectra of the successive elements starting with hydrogen as a means of deducing the quantum numbers of the inner electrons of the heavier elements. Errors which might otherwise creep in by this process may be eliminated by checking the first spark spectrum of each element against the normal arc spectrum of the element which precedes it in the periodic table.

Having justified the general procedure begun on pp. 482 to 483, we proceed with our discussion of the empirical evidence, turning next to the trivalent earth metals and to aluminum, in particular, since it follows

sodium and magnesium in the periodic table. Again we find two systems of energy levels—"doublet" and "quadruplet" this time. The old classification of the stationary states is applicable, but the normal state of the valence electron for aluminum is not 3s but 3p. Since the 3s level should have the lower energy if it existed—due to greater penetration into the core—we are driven to the conclusion that we have to do with a new exclusion rule. An examination of the spectrum of Al+ shows that the core of normal Al is similar to the normal state of Mg and contains two 3s electrons. The spectra of the other earth metals are similar and point to the general rule that *there can never be more than two electrons in s states of a given atom having the same total quantum number*. Mention should be made in this connection of the transition from helium, with two 1s electrons in the normal state, to Li with two 1s electrons and a 2s electron. Since there are no 1p states, the operation of the above rule in this case throws the valence electron of Li into a state of higher total quantum number than the inner electrons.

The spectra of the more electronegative elements which belong in the lower portions of the columns in the Thomsen diagram are much more complicated and for the present we omit all discussion of their analysis. Fortunately this analysis is unnecessary. There is a gradual, though somewhat irregular, increase in the *ionization potential* of the atoms from $Z = 5$ to $Z = 10$ and from $Z = 13$ to $Z = 18$, followed in each case by an abrupt fall from maxima at the inert gases neon and argon to minima at the succeeding alkali metals. Now, an increase of ionization potential should accompany the addition of more electrons to a group having a common total quantum number, for the increase in nuclear charge would not be wholly compensated by the additional repulsive potential of the group. On the other hand, a drop in ionization potential may be expected to occur when the normal state of the valence electron starts a group of higher total quantum number than any of the core electrons. Thus we are led to assume that the electrons added in transforming beryllium into neon go into 2p states, while those added in transforming magnesium into argon go into 3p states. 3d electrons are excluded from the latter group because the 3p electrons, on account of their greater penetration, are much more tightly bound.

From the fact that the inert gas neon is followed by the alkali sodium with a sudden drop in ionization potential we infer that the building up of the group of 2p electrons stops when the membership of that group reaches the value 6. Similarly the occurrence of the alkali potassium after the inert gas argon shows that the maximum number of 3p electrons is 6.

Without carrying farther this discussion of the empirical evidence for the formation of successive groups of electrons with higher and higher quantum numbers we can proceed to a provisional formulation of the Pauli principle. Our analysis shows that it takes two electrons to

complete a group of s electrons of given total quantum number, but six electrons for a group of p electrons of given total quantum number. This difference between the sizes of the complete s groups and the complete p groups becomes understandable when we remember that the p states are degenerate and include three substates each, corresponding to the values 0, ± 1 for the magnetic quantum number m. An obvious and satisfactory hypothesis is to assume that *there can be two, but only two, electrons in an atom having any given set of single-electron quantum numbers n, l, m*. This statement constitutes an elementary form of the *Pauli exclusion principle*.[1] It *excludes* the crowding of an indefinitely large number of electrons into any individual electron level in the limiting case of problem A and excludes all the energy levels of problem B correlated with excluded or overcrowded states of A. The principle has been substantiated by a mass of supporting evidence which would be out of place in this book. We refer the reader to the literature for further information.

In the form given above, the principle assumes our right to assign individual m values to the different electrons as we have assigned individual n and l values. Such an assignment cannot be justified by our previous argument. Nevertheless we can harmonize the rule with the wave-mechanics point of view if we modify its statement to read, *All energy levels of an atom are excluded whose zero-order approximations* (*problem A*) *involve products like* (57·6) *which assign more than two electrons to the same set of quantum numbers n, l, m.*

58. THE ELECTRON-SPIN HYPOTHESIS

58a. The Empirical Fine Structure of Spectrum Lines.—In its classical form the hypothesis of electron spin assumes that electrons are spherical shells of electricity in rotation about their own centers and consequently possess an angular momentum about this center and an intrinsic magnetic moment. The hypothesis was independently proposed by several different authors for various reasons, but first took definitive form and attracted wide attention when Uhlenbeck and Goudsmit[2] showed that it could be used to overcome the difficulties inherent in Landé's magnetic-core theory of the fine structure of spectrum lines.[3] In adapting the spin postulate to the principles of quantum

---

[1] W. PAULI, JR., *Zeits. f. Physik* **31**, 765 (1925); PAULING and GOUDSMIT, *loc. cit.*, Chap. IX.

[2] G. E. UHLENBECK and S. GOUDSMIT, *Die Naturwissenschaften* **13**, 953 (1925); *Nature* **117**, 264 (1926). The hypothesis was independently proposed by F. R. Bichowsky and H. C. Urey, *Proc. Nat. Acad. Sci.* **12**, 80 (1926). Earlier proponents of the hypothesis include A. H. Compton, *Wash. Acad. Sci. J.* **8**, 1 (1918), and E. H. Kennard, *Phys. Rev.* **19**, 420 (1922).

[3] A. LANDÉ, *Verh. d. deut. physik. Gesell.* **21**, 585 (1919); W. HEISENBERG, *Zeits. f. Physik* **8**, 273 (1922).

mechanics the model has been replaced by a system of equations which bear a close relation to the model but which cannot be derived from it in any rigorous sense. Heisenberg and Jordan[1] first treated the spin properties of the electron by the matrix method and later Pauli[2] gave a provisional formulation in terms of wave mechanics. A very great advance was made when Dirac[3] showed that a more rigorous wave-mechanical treatment can be derived as a by-product of an attempt to reconcile the quantum theory with special relativity theory. Dirac's method of approach dispenses entirely with all assistance from the classical model of Uhlenbeck and Goudsmit and is now generally conceded to be the most fundamental way to introduce the subject. It is relatively abstract, however, and leads ultimately to great difficulties which have not been fully overcome as yet.[4] Since the Pauli spin theory is easier to apply than Dirac's, and yet sufficiently accurate for most purposes, it seems wise to give this elementary formulation here, leading up to it by the historical-empirical approach. We begin with a brief description of the fine structure of optical energy-level systems.[5]

The optical energy-level spectra of the alkaline-earth metals show an empirical structure similar to that which we have found in the case of the alkalies, but more complicated, inasmuch as there are two sharp series, two principal series, etc. In fact we find two complete arrays of energy levels, each of which, fine structure excepted, is a replica of the energy-level array of one of the alkalies. Transitions between one major level, *i.e.*, one group of fine-structure components, and another of the same array are governed by the selection rule $\Delta l = \pm 1$. Transitions between levels of the different arrays occur much less freely than those between levels of the same array, especially in the spectra of the lighter elements of this type. One of these arrays (singlet) shows no fine structure, but, outside the sharp series, whose members are single, the levels of the other array (triplet) have three closely spaced components each.

All energy levels which can be given a place in such an array can be accounted for in first approximation by means of a single-electron model. Not *all* major energy levels of the alkaline-earth elements, however, can

[1] W. HEISENBERG and P. JORDAN, *Zeits. f. Physik* **37**, 263 (1926).

[2] W. PAULI, JR., *Zeits. f. Physik* **43**, 601 (1927).

[3] P. A. M. DIRAC, *Proc. Roy. Soc.* **A117**, 610, **A118**, 351 (1928). *Cf.* also Dirac *P.Q.M.*, Chap. 13; C. G. Darwin, *Proc. Roy. Soc.* **A118**, 654 (1928).

[4] *Cf.* W. PAULI, in Geiger and Scheel's *Handbuch der Physik*, 2d ed., vol. **24**, Part I, pp. 242–247 (1933); also W. H. FURRY and J. R. OPPENHEIMER, *Phys. Rev.* **45**, 245 (1933).

[5] Here we refer to the coarser type of fine structure which originates in the spin of the extranuclear electrons. There is an additional so-called "hyperfine structure" due to interaction between the outer electrons and the spin of the nucleus concerning which we shall have nothing to say.

be fitted into a scheme of classification based on assignments of a quantum number l which conforms to the principle of selection $\Delta l = \pm 1$. Moreover, although structure of this type is to be found in the energy-level spectra of elements with more than two valence electrons, its incompleteness becomes more evident as the valence increases. On the other hand it is possible to include many additional energy levels in a classification of the major levels which assigns to each an l value in such fashion as to preserve the selection principle $\Delta l = 0, \pm 1$ and to divide the levels of most atoms into *systems* analogous to the singlet and triplet arrays described above and each having its own characteristic fine structure.

The argument of Sec. 40e indicates that the square of the resultant angular momentum ($\mathfrak{L}^2$) should be an integral of the motion of a free atomic system composed of point electrons and a point nucleus. Furthermore, $\mathfrak{L}^2$ should be a function of H so that each energy level is characterized by a definite value of the angular-momentum quantum number L. According to the Bohr theory radiative transitions between states of a many-particle problem with different angular momenta should be governed by the principle of selection $\Delta L = 0, \pm 1$. The same rule will be derived on a quantum-mechanical basis in Sec. 64. Since this selection rule is broader in scope than the rule for one-electron systems, we conclude that the empirical quantum number which changes by $0, \pm 1$ is to be identified with the quantum number L of the resultant angular momentum of all the electrons. From this point on we reserve the letters s, p, d, f, $\cdots$, introduced on p. 475 for the designation of the states of a single electron moving in a central force field, and introduce capital letters for states of a complete atom having different angular momenta. The correlation between the letters and the L values is indicated in Table I, p. 494.

All the major energy levels with a given value of L and belonging to a common system have the same number of fine-structure components. Within any system the number of such components increases in steps of one or two as we pass from one value of L to the next larger one until a maximum value is reached, after which the number of sublevels remains constant. This maximum multiplicity is the fundamental characteristic of the system, which is called "singlet," "doublet," "triplet," "quadruplet," etc., according as its value is $1, 2, 3, 4, \cdots$. Table I shows in detail the number of components for the major levels of all types in systems with multiplicity up to 4.

Space does not permit a discussion of the methods by which empirical spectroscopists have been able to unravel these structures from the maze of lines which make up the spectrum from which they start. Suffice it to note here that the relative weakness of intersystem combination lines and the characteristic Zeeman patterns of different types of line are important aids.

TABLE I.—MULTIPLICITIES OF DIFFERENT CLASSES OF ENERGY LEVELS

Symbol	S	P	D	F	G	H	...
L	0	1	2	3	4	5	...
Singlet system....................	1	1	1	1	1	1	...
Doublet system...................	1	2	2	2	2	2	...
Triplet system....................	1	3	3	3	3	3	...
Quadruplet system...............	1	3	4	4	4	4	...

It will be observed that the number of components is always equal to $2L + 1$ unless that number exceeds the multiplicity of the system. This rule is confirmed by the study of systems of still higher multiplicity. The number $2L + 1$ is equal to the number of $\mathcal{L}_z$ eigenvalues consistent with the eigenvalue $L(L + 1)h^2/4\pi^2$ for $\mathcal{L}^2$, i.e., to the number of sublevels into which a major level of angular-momentum quantum number L should be split by a magnetic field according to the elementary spin-free theory of the Zeeman effect given in Sec. 49c. In all cases the spacing of the fine-structure components decreases rapidly with increasing values of L and of the ordinal number in a Rydberg series.

The fine structure of the energy-level system is, of course, derived from the corresponding fine structure of the spectrum lines. The latter also reveals the existence of a new principle of selection affecting transitions between different sublevels of any pair of major levels. This rule takes the simple form $\Delta J = 0$, ± 1 if we assign appropriate values of a new quantum number J, called the *inner quantum number*. A suitable scheme for the assignment of J values, which gives the correct number of components for each major level, is contained in the following rule.

Let $2S + 1$ denote the multiplicity of the system to which the given major level belongs. The maximum value of J for any of the components is $L + S$ and the minimum value $|L - S|$. The remaining values are spaced at equal unit intervals between these extremes. Consider, for example, the case of a P level in the quadruplet system. Here $S = \frac{3}{2}$ and $L = 1$. The J values are $2\frac{1}{2}$, $1\frac{1}{2}$, $\frac{1}{2}$. A D level of the same system will have four components for which J takes on the values $3\frac{1}{2}$, $2\frac{1}{2}$, $1\frac{1}{2}$, $\frac{1}{2}$, respectively. The so-called "normal" order of the energy values is that of increasing J and is the usual order for the simple spectra of the electropositive elements. It is more or less systematically inverted in the case of the electronegative elements.

It is evident at once that the above normal assignment of J values is arbitrary in that we could increase all the values in the spectrum of any element by any amount J_0 without affecting the form of the selection principle. According to the normal assignment the J values are integral in the case of systems of odd multiplicity and half-integral in the case of systems having even multiplicity. It is a fundamental and very

important empirical rule that *systems of even multiplicity only occur in atoms or ions having an odd number of extranuclear electrons, while systems of odd multiplicity only occur in atoms or ions with an even number of extranuclear electrons.* Hence it would be possible to eliminate the half-integral J values entirely by introducing a J_0 which has a half-integral value for atoms and ions with an odd number of extranuclear electrons, and an integral value for those which have an even number of extranuclear electrons. Such an arbitrary procedure would partially spoil the symmetry of the normal assignment of J values, however, and does not lend itself to the development of the theory.

58b. The Combination of Angular Momenta.[1]—The existence of *any* fine structure in the spectra of the alkali atoms means that the single-electron model is imperfect, and that, if it is to be retained at all, the central-force-field hypothesis must be modified. Classically such a modification would imply that the optical electron is subject to a torque with respect to the nucleus as a center. In the case of a free atom such a torque would imply a countertorque and hence a second angular momentum in the atom. The natural inference would be that the second angular momentum comes from the core electrons. Of course the coupling between the optical electron and the core electrons would have to be weak to account for the fact that the single-electron model works so well. Since the principle of selection for the inner quantum number J is the same as that to be expected for the quantum number of the square of the resultant angular momentum, we are led to identify J with that quantum number and l, or L, with the quantum number of the square of the angular momentum of the optical electron, or the optical group of electrons, as the case may be. In the case of S states the angular momentum of the optical electron, or electrons, is, on this hypothesis, zero. According to the normal assignment the corresponding value of J is the quantum number S which fixes the multiplicity of the system. We are thus led to identify the square of the core angular momentum with $S(S + 1)h^2/4\pi^2$.

Let us now examine the number of fine-structure components of the various major energy levels predicted by this provisional theory. To this end we consider the coupling of two sets of electrons revolving about a common nucleus. To avoid prejudice and confusion we introduce new symbols for the quantum numbers in this calculation. Let L_1 and L_2 denote the angular-momentum quantum numbers for the two individual sets, while L denotes the quantum number of the resultant angular momentum. The corresponding magnetic quantum numbers will be indicated by M_1, M_2, and M, respectively. We refer to the uncoupled state of the system, in which the interaction terms of the Hamiltonian are neglected or replaced by centrally symmetric averages, as case A,

[1] The method here used is due to Slater [*Phys. Rev* **34**, 1293 (1929)].

and to the coupled state, as case B. For case A one can use a system of wave functions which quantize the square of the total angular momentum of each group and the z component of the angular momentum of each group. The matrices of $\mathcal{L}_1{}^2$, $\mathcal{L}_2{}^2$, $\mathcal{L}_{1z}$, and $\mathcal{L}_{2z}$ are then diagonal. We can also use a system of wave functions which quantizes the square of the resultant angular momentum of the combination of the two groups and the z component of that momentum. The latter system of wave functions is adapted to the description of the coupled state, case B, in which $\mathcal{L}_1{}^2$, $\mathcal{L}_2{}^2$, $\mathcal{L}_{1z}$, $\mathcal{L}_{2z}$ are no longer integrals of the motion, but the former description is the one which gives meaning to the assignment of separate quantum numbers to the two groups.

Let us now consider a case A energy level with the individual quantum numbers L_1 and L_2. Ignoring permutation degeneracy—this procedure will be justified later on—we note that the level in question has

$$(2L_1 + 1)(2L_2 + 1)$$

linearly independent wave functions obtained by multiplying the wave function of the first group for each value of M_1 by the wave function of the second group for each value of M_2. A suitable canonical transformation will carry these wave functions over into a new set, equal in number, which are eigenfunctions of the angular momenta $\mathcal{L}^2$ and $\mathcal{L}_z$ of the combination. As a matter of fact each of the product functions is *already* an eigenfunction of $\mathcal{L}_z$ with the eigenvalue $M = M_1 + M_2$. Hence the canonical transformation will form linear combinations of those initial functions only which have a common value of $M_1 + M_2$. We are thus led to construct a diagram of the type shown in Table II which illustrates the special case $L_1 = 2$; $L_2 = 1$. The M_1 values are

TABLE II.—VALUES OF $M_1 + M_2$ FOR CASE WHERE $L_1 = 2$; $L_2 = 1$.

		M_1				
		-2	-1	0	$+1$	$+2$
M_2	-1	-3	-2	-1	0	$+1$
	0	-2	-1	0	$+1$	$+2$
	$+1$	-1	0	$+1$	$+2$	$+3$

laid off horizontally, the M_2 values vertically, each square indicating a possible product function for case A. In these squares we have inserted the corresponding values of M. In this particular case we shall have one product function for $M = \pm 3$; two for $M = \pm 2$; three for $M = 0, \pm 1$. In passing from the case A product functions to the zero-order case B functions we have to "scramble" the different product functions belonging to each M value. As a result we obtain the same number of case B

functions for each M value that we had before. Thus each square of Table II may be considered as representing a case B function with the given value of M.

Knowing the number of substates for each M value in case B we can readily determine the resulting eigenvalues of $\mathcal{L}^2$, or the values of the quantum number L. Since 3 is the largest value of M in the group represented by the table, it follows from the discussion in Sec. 40f that the states for which $M = +3$ must have the quantum number $L = 3$. But we know from this same discussion that if an energy level of a free atomic system has any eigenfunctions which are also eigenfunctions of $\mathcal{L}^2$ with the eigenvalue $L(L + 1)h^2/4\pi^2$, it must have $2L + 1$ such functions with the "magnetic" quantum numbers $M = 0, \pm 1, \pm 2, \cdots, \pm L$. The construction of such an L complex will use up all the case B functions on the top row and right-hand column of the array shown in Table II. There will then be just one remaining case B function for each of the M values, ± 2. Consequently we must have an L complex for $L = 2$ as well as $L = 3$. This will use the remainder of the second row of the figure and one function from the third row. Finally the remaining functions in the last row are just sufficient to form an L complex with the eigenvalue $L = 1$.

In general it will be evident that the number of L complexes which are formed by a canonical transformation of the $(2L_1 + 1)(2L_2 + 1)$ product functions of an (L_1, L_2) case A energy level is equal to the number of rows in a diagram like Table II arranged with the long way of the diagram horizontal. In other words it is equal to the smaller of the two quantities $(2L_1 + 1)$, $(2L_2 + 1)$.

The application of the perturbing Hamiltonian which transforms the case A problem into the case B problem cannot break up any of the L complexes since, by the theory of Sec. 40f, they are indissoluble so long as there is no external torque applied to the system as a whole. On the other hand, since this perturbation brings in Coulomb forces between individual electrons of the two different sets, it must eliminate the degeneracy with respect to $\mathcal{L}_{1z}$ and $\mathcal{L}_{2z}$, thus producing energy differences between the different L complexes. From the standpoint of Sec. 40e this splitting of the unperturbed energy levels is to be attributed to the fact that the perturbing term of the Hamiltonian does not commute with the angular-momentum vectors of the two parts of the system. Hence there remain no operators which commute with H which do not also commute with $\mathcal{L}^2$. This makes $\mathcal{L}^2$ a function of H (except for eventual accidental and continuous-spectrum degeneracy) and allows only one value of L for any one energy level. We conclude that a case A energy level characterized by the quantum numbers L_1, L_2 should be split by the interaction into $2L_1 + 1$ components, where L_1 is taken to be the smaller of the two numbers L_1, L_2.

A less rigorous, but perhaps more appealing, argument for this splitting is the following. Since each of the zero-order case B wave functions is a simultaneous eigenfunction of $\mathcal{L}^2$, $\mathcal{L}_1^2$, $\mathcal{L}_2^2$, the relation

$$\mathcal{L}^2 = \mathcal{L}_1^2 + \mathcal{L}_2^2 + 2\vec{\mathcal{L}}_1 \cdot \vec{\mathcal{L}}_2 \tag{58·1}$$

demands that two such eigenfunctions belonging to different L complexes must be eigenfunctions of

$$\vec{\mathcal{L}}_1 \cdot \vec{\mathcal{L}}_2 \equiv \mathcal{L}_{1x}\mathcal{L}_{2x} + \mathcal{L}_{1y}\mathcal{L}_{2y} + \mathcal{L}_{1z}\mathcal{L}_{2z}$$

with different eigenvalues. Classically the energy could not be independent of $\vec{\mathcal{L}}_1 \cdot \vec{\mathcal{L}}_2$ unless the mutual energy of the two groups of particles was independent of their orientation. But if the two groups have nonvanishing angular momenta they cannot have spherical symmetry, and their mutual energy cannot be independent of their orientation. Hence we infer that the average energy of zero-order case B eigenfunctions belonging to different L complexes must be different.

58c. The Landé Magnetic Core Theory.—Turning back to the problem of the actual fine structure of spectrum lines, we are led to correlate the quantum numbers L, L_1, L_2 of the preceding theory with the empirical quantum numbers J, L, S, respectively, of Sec. 58a. In other words we postulate that the atom can be resolved into two weakly coupled parts with individual angular-momentum vectors $\vec{\mathcal{L}}$ and $\vec{S}$, whose resultant $\vec{\mathcal{L}} + \vec{S}$ we shall call $\vec{\mathcal{J}}$. In applying the term "angular momentum" to these vectors we imply that they are subject to the commutation rules (38·15), (38·16), and (38·17). It follows from Sec. 40f that the associated eigenvalues are given by expressions of the form

$$\left.\begin{array}{ll}
(\mathcal{L}^2)' = L(L+1)\dfrac{h^2}{4\pi^2}, & (\mathcal{L}_z)' = M_L\dfrac{h}{2\pi}, \\[2mm]
(S^2)' = S(S+1)\dfrac{h^2}{4\pi^2}, & (S_z)' = M_S\dfrac{h}{2\pi}, \\[2mm]
(\mathcal{J}^2)' = J(J+1)\dfrac{h^2}{4\pi^2}, & (\mathcal{J}_z)' = (\mathcal{L}_z + \mathcal{J}_z)' = M\dfrac{h}{2\pi}.
\end{array}\right\} \tag{58·2}$$

The interpretation of $\vec{\mathcal{L}}$ and $\vec{S}$ as ordinary orbital angular momenta defined by (34·10) requires that the quantum numbers L, M_L, S, M_S, J, M shall all be integers, but the general theory of Sec. 40f admits the possibility of odd multiples of $\frac{1}{2}$. According to the above hypothesis the number of sublevels due to a weak coupling of these vectors should

be $2L + 1$ if $L \leqslant S$, and $2S + 1$ if $S \leqslant L$—a result in agreement with observation.[1]

Interpreting $\vec{\mathcal{L}}$ as the angular momentum of the exterior (valence) electron or electrons active in producing the optical spectrum, we are confronted with the problem of finding a physical interpretation for $\vec{s}$. The first and obvious suggestion that $\vec{s}$ is the orbital angular momentum of the core electrons came from Landé. He observed that the existence of a core angular momentum implied the existence of a corresponding core magnetic moment whose field would apply the necessary torque to the valence electron. In first approximation the interaction energy of the two angular momenta was taken to be the negative scalar product of the magnetic moment of the core and the average magnetic field due to the valence electron at the nucleus. The absolute value of the width of the alkali doublets calculated on this basis is of the right order of magnitude, but the Landé theory involved several serious difficulties of which three will be mentioned here.

One of these difficulties was the inversion of the energetic order of the fine-structure levels in normal multiplets. According to the Landé theory a parallel orientation of the two angular momenta should give the least energy and, of course, the largest J value. The opposite energetic order is the normal one, empirically. A second source of embarrassment lay in the occurrence of half-integral values of S and J for elements of odd atomic number—an empirical feature equally incomprehensible from the standpoint of the Bohr theory and of the wave mechanics, so long as one uses the Landé model. The third and crucial difficulty lay in the existence of a discrepancy between the angular momentum of the normal state of each atom and that of the core of the succeeding element of the periodic table. The inert gases are all diamagnetic and we should accordingly expect the normal state of each of them to have zero magnetic moment and zero angular momentum. Thus the J value of the normal state of neon and the S value of the sodium atom should both be zero. Actually we find that the latter is $\frac{1}{2}$. Again, the J value of the normal state of sodium being $\frac{1}{2}$, we should expect the S value of magnesium to be the same. Actually there are two systems of series in the magnesium energy-level scheme, one singlet ($S = 0$), the

---

[1] It is worthy of note that the same result is obtained if one asks for the number of different orientations of a vector of length L with respect to a vector of length S which makes the length of their resultant an integer or a half integer, according as the sum $L + S$ is assumed to be an integer or a half integer. In the Bohr theory L and S were interpreted directly as the measures of the absolute values of the corresponding angular momenta. Thus a simple vector diagram gave the result regarding the interaction of angular momenta which we have derived somewhat laboriously above.

other triplet ($S = 1$). It is not possible to "doctor" the J value assignments so as to eliminate these discrepancies.

58d. Solution of the Fine-structure Problem by the Electron-spin Hypothesis.—The last mentioned discrepancy in the theory suggested to Uhlenbeck and Goudsmit[1] the possibility that the electrons have an angular momentum due to rotation about an internal axis—"electron spin"—in addition to their orbital angular momentum. On this hypothesis the fine structure of the spectra of sodium and the other alkalies is interpreted as due to the interaction between the orbital and spin angular momenta of the valence electron alone, the core having zero angular momentum and playing only a minor passive part in the problem. Since all angular momenta which need be considered belong to a single electron, we shall use the small letters l, s, j, m_l, m_s, m, instead of L, S, J, M_L, M_S, M, for their quantum numbers. In order to account for the doublet fine structure we have only to identify the value $S = s = \frac{1}{2}$ characteristic of sodium with the quantum number which determines the resultant electron-spin angular momentum. In other words we assume that the square of the spin angular momentum has the single eigenvalue $\frac{1}{2}(\frac{1}{2} + 1)h^2/4\pi^2$. Assuming further that the spin angular momentum combines with other angular momenta just as two orbital angular momenta combine, we infer that the j values of the resultant of the spin and orbital angular momenta should be $l + \frac{1}{2}$ and $l - \frac{1}{2}$ provided that $l > 0$. If $l = 0$, j reduces to the spin quantum number $\frac{1}{2}$. This checks with the empirical results for sodium and the other alkalies.

In the case of magnesium and the other alkaline earths, the spins of the two valence electrons must be assumed to combine first with each other to form a resultant spin angular momentum with quantum numbers 0 or 1. The observed fine structure is then interpreted as due to the combination of the orbital angular momentum of the optical electron with the resultant spin angular momentum.

Classical theory requires that the spin angular momentum shall be accompanied by a corresponding spin magnetic moment. The ratio of the magnetic moment of a spherical spinning electron to the corresponding angular momentum had been calculated before the promulgation of the Uhlenbeck and Goudsmit hypothesis by Abraham, who found it to be $e/\mu c$, or *twice* the value of the same ratio for the orbital motion of an electron. (The formula is algebraic: the negative charge e goes with a magnetic moment opposite in direction to the spin angular momentum.) Hence the spin axis should precess about the lines of an external magnetic field with a double-normal Larmor precession frequency.

If the spin angular momentum of an electron is free to orient itself in any direction, and is subjected to the influence of an external magnetic field, we should expect the component of that momentum in the direction

[1] *Loc. cit.*, footnote 2, p. 491.

of the field, say S_z, to be quantized just as the component of the orbital angular momentum parallel to the field is quantized in the elementary theory of the Zeeman effect. The eigenvalues of S_z compatible with the resultant spin angular momentum $(S^2)' = \frac{1}{2}(\frac{1}{2} + 1)h^2/4\pi^2$ are evidently $\pm\frac{h}{4\pi}$. Classically the mutual energy of the field and spin would be equal to the negative scalar product of the magnetic moment and the field strength $\overrightarrow{\mathcal{H}}$. The eigenvalues of the mutual energy are accordingly $\pm\frac{e\mathcal{H}h}{4\pi\mu c}$, where $\mathcal{H} = |\overrightarrow{\mathcal{H}}|$. In the sharp-series energy levels of sodium the electron spin is uncoupled and we should accordingly expect these levels to be split in an external magnetic field into two components symmetrically placed with respect to the unperturbed line and having the energy difference $e\mathcal{H}h/2\pi\mu c$, i.e., twice the spacing of the magnetic sublevels in the simple Zeeman effect of Sec. 49c. An empirical analysis of the complex Zeeman effect for sodium reveals the expected splitting of the sharp-series energy levels and thus checks the assumed ratio of the magnetic moment to the spin angular momentum, as well as our general procedure. A very important additional check comes from the measurements of the so-called "gyromagnetic ratio" in ferromagnetic solids.[1]

We have next to examine the nature of the interaction between the spin angular momentum and the orbital angular momentum of an electron. A proper treatment of this interaction yields a quantitative explanation of the fine structure of the alkali-metal spectra and the basis for a theory of the fine structure of the hydrogen-atom spectrum. The existence of *any* interaction implies a torque exerted on the spin due to the motion of the electron through the electric field of the nucleus and a countertorque about the nucleus as a center and tending to change the nuclear angular momentum. The reason for these torques becomes evident when we recollect that an electrostatic field can be transformed into an electromagnetic one by the application of a Lorentz transformation. In a frame of reference in which the electron is momentarily at rest and the nucleus moving with a velocity $\overrightarrow{v}$ the electron is subject to a magnetic field

$$\overrightarrow{\mathcal{H}_n} = \frac{Z|e|}{c}\frac{\overrightarrow{r} \times \overrightarrow{v}}{r^3},$$

where $\overrightarrow{r}$ is the vector distance from the electron to the nucleus. The electron in this field will have a classical energy

[1] *Cf.* S. J. BARNETT, *Rev. Mod. Phys.* **7**, 129 (1935).

$$U_s' = \frac{e}{\mu c}\vec{s}\cdot\vec{\mathcal{K}}_n = \frac{Ze^2}{\mu c^2}\left(\frac{1}{r^3}\right)\vec{s}\cdot[\vec{r}\times\vec{v}]$$

due to its spin.   If we suppose this energy to be independent of the frame of reference and note that $\mu\vec{r}\times\vec{v}$ is equal to the orbital angular momentum in that frame of reference in which the nucleus is at rest, we obtain the classical expression

$$U_s' = Ze^2\frac{\vec{s}\cdot\vec{\mathcal{L}}}{\mu^2 c^2 r^3} \tag{58.3}$$

for the mutual energy of the two angular momenta.

A more careful relativistic analysis shows, however, that the above expression should properly be multiplied by $\frac{1}{2}$—the famous Thomas correction factor.[1]   Of the various discussions of this factor which have been published, that by Kramers is perhaps the most luminous.   He shows that if the torque applied to the spin angular momentum of an electron at rest in a magnetic field is $(e/\mu c)\vec{s}\times\vec{\mathcal{K}}$ the corresponding torque for an electron moving with velocity $\vec{v}$ is given by the equation

$$\frac{d\vec{s}}{d\tau} = \frac{e}{\mu c}\left\{\vec{s}\times\vec{\mathcal{K}} + \frac{1}{2c}\vec{s}\times[\vec{\mathcal{E}}\times\vec{v}]\right\} + \vec{T}, \tag{58.4}$$

in which τ is the proper time, and $\vec{T}$ is defined by

$$\vec{T} = \frac{e}{\mu c^2}\left\{\vec{s}(\vec{v}\cdot\vec{\mathcal{E}}) - \frac{1}{2}\vec{\mathcal{E}}(\vec{v}\cdot\vec{s}) - \frac{1}{2}\vec{v}(\vec{\mathcal{E}}\cdot\vec{s})\right\}.$$

If the electron is in not too rapid motion about a center of force, we can replace the electric force $\vec{\mathcal{E}}$ by $(\mu/e)\,d\vec{v}/d\tau$ and thereby reduce $\vec{T}$ to the

---

[1] L. H. THOMAS, *Nature* **117**, 514 (1926), *Phil. Mag.* **3**, 1 (1927); J. FRENKEL, *Zeits. f. Physik* **37**, 243 (1926); H. A. KRAMERS, *Physica* **1**, 825 (1934).

Recent letters to the editor of the *Physical Review* by Inglis and by Dancoff and Inglis [*Phys. Rev.* **50**, 783, 784 (1936)] bring out an important point regarding the Thomas correction which is somewhat obscured from the dynamical point of view of Kramers.   The corrected spin-orbit interaction energy U_S of Eq. (58.6) is shown to be the sum of the "magnetic" energy term U_S' of Eq. (58.3) and a negative kinematic term of relativistic origin which has the form

$$U_S'' = -\frac{1}{2c}\left(\vec{s}\cdot\left[\vec{v}\times\frac{d\vec{v}}{dt}\right]\right).$$

U_S'' is equal to $-\frac{1}{2}U_S'$ in the case of an electron in orbital motion in a Coulomb field. On the other hand in atomic nuclei, forces of a nonelectric character predominate and U_S'' is greater in absolute value than U_S'.   This causes a reversal of the sign of the spin-orbit interaction for nuclei.

form $\dfrac{1}{2c^2}\dfrac{d}{d\tau}\{\vec{v}\times[\vec{s}\times\vec{v}]\}$. For motion in a closed orbit $\vec{T}$ would average

to zero. The secular changes in $\vec{s}$ are then given by (58·4) with the

term $\vec{T}$ omitted. The average torque for the time of one revolution is

thus perpendicular to $\vec{s}$ and the secular changes alter its direction but

not its magnitude. The corresponding energy for an electron in orbital

motion through an electrostatic central force field with potential Φ is

$$U_S = -\frac{e\vec{s}\cdot[\vec{\mathcal{E}}\times\vec{v}]}{2\mu c^2} = \frac{e}{2\mu^2 c^2}\left(\frac{1}{r}\frac{d\Phi}{dr}\right)\vec{s}\cdot\vec{\mathcal{L}} = \frac{1}{2\mu^2 c^2}\left(\frac{1}{r}\frac{dV}{dr}\right)\vec{s}\cdot\vec{\mathcal{L}}. \quad (58\cdot5)$$

In the case of a Coulomb field with charge $-Ze$ at the center of force the
expression for the energy reduces to

$$U_S = \frac{Ze^2}{2\mu^2 c^2}\frac{\vec{s}\cdot\vec{\mathcal{L}}}{r^3} \quad (58\cdot6)$$

or half the value given by (58·3). Except for the difference in sign
already noted, this final classical formula is similar in form to the energy

expression used by Landé, which also involves the factor $\dfrac{e^2}{\mu^2 c^2}\dfrac{\vec{s}\cdot\vec{\mathcal{L}}}{r^3}$ when

properly reduced, but not the factor Z. It will be observed that (58·6)

makes the energy greatest for parallel orientations of $\vec{s}$ and $\vec{\mathcal{L}}$, so that
it has the correct sign and eliminates the first of the difficulties in the
Landé theory mentioned above.

59. THE FINE STRUCTURE OF THE SPECTRA OF ATOMIC SYSTEMS WITH A SINGLE VALENCE ELECTRON

The work of Sec. 58 gives the electron-spin hypothesis fairly definite
form and indicates a decided superiority over the Landé magnetic-core
theory as an explanation of the fine structure of atomic spectra. We
proceed to the crucial test of a quantitative calculation of the spacing
of the fine-structure levels in the spectra of the alkalies and of hydrogen.
Incidentally we shall derive the formulas appropriate to positive ions
isoelectronic to one or the other of the above types of atoms.

Let us begin with a first-order approximation to the energy correction
for the spin-orbit interaction in hydrogen and the hydrogenic states of
the alkalies. For this purpose we need to compute the mean value of U_S
for the zero-order case B wave functions of Sec. 58*b* or, what amounts to
the same thing, the diagonal terms of the matrix of U_S in a scheme of
wave functions which makes the unperturbed energy, the square of the
orbital angular momentum, that of the spin angular momentum, and
that of the resultant angular momentum, all diagonal. We reserve for
later discussion the question of the exact nature of the wave functions

to be used in the study of electron spin and proceed on the assumption that from a matrix point of view the interaction between the spin angular momentum and the orbital angular momentum can be treated in the same way as that of two orbital angular momenta.

Let $\vec{L}, \vec{S}, \vec{J}$ denote respectively the matrices of $\vec{\mathcal{L}}, \vec{\mathcal{S}},$ and $\vec{\mathcal{J}}$. We know from Sec. 58b that when L^2, $L_1{}^2$, and $L_2{}^2$ are diagonal, $\vec{L}_1 \cdot \vec{L}_2$ is diagonal. Similarly, when L^2, S^2 and

$$J^2 = (L_x + S_x)^2 + (L_y + S_y)^2 + (L_z + S_z)^2 \qquad (59 \cdot 1)$$

are diagonal, the scalar product $\vec{S} \cdot \vec{L}$ must be diagonal. In fact (59·1) leads directly to the relation [*cf.* (58·1)]

$$J^2 = L^2 + S^2 + 2\vec{S} \cdot \vec{L}, \qquad (59 \cdot 2)$$

if we assume that the components of $\vec{L}$ commute with those of $\vec{S}$. From (59·2) we can compute the eigenvalues of $\vec{S} \cdot \vec{L}$. Thus, using the quantum numbers of $\mathcal{S}^2$ and $\mathcal{L}^2$ as matrix indices and suppressing the other indices required for a complete notation, we have (*cf.* p. 366)

$$(j,l|\vec{S} \cdot \vec{\mathcal{L}}|j,l) = \frac{h}{8\pi^2}\left[j(j+1) - l(l+1) - \frac{1}{2}\left(\frac{1}{2}+1\right)\right]. \quad (59 \cdot 3)$$

The possible values of j are $l + \frac{1}{2}$ and $l - \frac{1}{2}$. Hence the two eigenvalues

$$\left. \begin{aligned} (l + \tfrac{1}{2},l|\vec{S} \cdot \vec{\mathcal{L}}|l + \tfrac{1}{2},l) &= \frac{lh^2}{8\pi^2}, \\ (l - \tfrac{1}{2},l)|\vec{S} \cdot \vec{\mathcal{L}}|l - \tfrac{1}{2},l) &= -\frac{(l+1)h^2}{8\pi^2} \end{aligned} \right\} \qquad (59 \cdot 4)$$

are obtained.

Equation (58·6) shows that the matrix $\mathbf{U}_S$ is a multiple of the product of $\vec{S} \cdot \vec{L}$ into the matrix of $1/r^3$. Since the former matrix is diagonal, the diagonal elements of $\mathbf{U}_S$ which give the first-order energy corrections are multiples of products of the eigenvalues of $\vec{S} \cdot \vec{\mathcal{L}}$ into diagonal elements of the matrix of $1/r^3$, *i.e.*, into the average values of $1/r^3$ for the appropriate orbital wave functions.

We proceed to a calculation of $\overline{1/r^3}$ for a state with total quantum number n and azimuthal quantum number l. As the radial wave functions are independent of the magnetic quantum number m_l, the mean values of $1/r^3$ are also independent of m_l. This is fortunate as the m_l values have been scrambled in forming the zero-order case B functions.

The calculation of $\overline{1/r^3}$ for hydrogenic states was first carried through on the basis of the Bohr theory.[1]   A wave-mechanical evaluation of this quantity has been worked out by Waller[2] using a somewhat laborious direct evaluation of the integrals $\int \mathfrak{R}_{nl}(r)^2 r^{-3} dr$.   The following discussion, however, is an adaptation of an ingenious matrix treatment due to Heisenberg and Jordan.[3]   The procedure involves two steps of which the first is the evaluation of $\overline{1/r^3}$ in terms of $\overline{1/r^2}$.

As a starting point we choose the scheme of wave functions

$$\chi_{n,l,m_l}(r,\theta,\varphi) = \mathfrak{R}_{nl}(r)\Theta_{lm_l}(\theta)\Phi_{m_l}(\varphi),$$

introduced in Sec. 28 and based on the use of $\int |\chi_{nlm_l}|^2 \sin\theta\, dr d\theta d\varphi$ as a normalizing integral.   From the corresponding radial equation (28·19) we see that the Hamiltonian operator appropriate to this type of wave function for the unperturbed problem, with $\vec{\mathcal{L}}$ and $\vec{S}$ uncoupled, is

$$H_a \equiv -\frac{h^2}{8\pi^2\mu}\frac{\partial^2}{\partial r^2} + \frac{\mathcal{L}^2}{2\mu r^2} - \frac{Ze^2}{r}, \tag{59·5}$$

where $\mathcal{L}^2$ is to be given the form of Eq. (34·18).

Denoting the Hermitian operator $\dfrac{h}{2\pi i}\dfrac{\partial}{\partial r}$ by p_r, let us consider the Poisson bracket

$$-[H_a,p_r] = \left[H_a\frac{\partial}{\partial r} - \frac{\partial}{\partial r}H_a \right] = \left[\frac{\mathcal{L}^2}{\mu r^3} - \frac{Ze^2}{r^2} \right]. \tag{59.6}$$

It follows from the fundamental formula (38·6) that the mean value of this operator for any state is the time rate of change of the mean value of p_r for that state.   In the case of a stationary state the mean value of p_r must be constant and that of $[H_a,p_r]$ must be zero.   If the state in question is an eigenstate for $\mathcal{L}^2$ we have

$$\int \chi_{nlm_l}{}^*\left(\frac{\mathcal{L}^2}{\mu r^3} - \frac{Ze^2}{r}\right)\chi_{nlm_l} \sin\theta\, dr d\theta d\varphi = \frac{l(l+1)h^2}{4\pi^2\mu}\left(\overline{\frac{1}{r^3}}\right) - Ze^2\left(\overline{\frac{1}{r^2}}\right) = 0.$$

Thus[4]

$$\left(\overline{\frac{1}{r^3}}\right) = \frac{4\pi^2\mu Ze^2}{h^2 l(l+1)}\left(\overline{\frac{1}{r^2}}\right). \tag{59·7}$$

[1] *Cf.* M. BORN, *Atommechanik*, §22.

[2] WALLER, *Zeits. f. Physik* **38**, 635 (1926).

[3] W. HEISENBERG and P. JORDAN, *Zeits. f. Physik* **37**, 263 (1926).   A more general matrix method for calculating the mean values of negative powers of r has been given by Van Vleck, *Proc. Roy. Soc.* A **143**, 679 (1934).   The derivation of (59·12) below also follows a procedure due to Van Vleck, *Quantum Principles and Line Spectra*, p. 299, 1926.

[4] A similar examination of the mean value of $[H,rp_r]$ for an arbitrary central force field yields

The eigenvalues of the Hamiltonian of Eq. (59 5) are those of the radial equation

$$H_a \Re(r) \equiv \left\{ \frac{1}{2\mu} \left[p_r{}^2 + \frac{l(l+1)h^2}{4\pi^2 r^2} \right] - \frac{Ze^2}{r} \right\} \Re(r) = E\Re(r), \quad (59\cdot8)$$

viz.,

$$E_n = -\frac{NhcZ^2}{n^2}.$$

To evaluate $\overline{1/r^2}$ we introduce the modified Hamiltonian

$$H_b = \frac{1}{2\mu} \left[p_r{}^2 + \frac{(l+\epsilon)(l+\epsilon+1)h^2}{4\pi^2 r^2} \right] - \frac{Ze^2}{r}.$$

The corresponding radial equation is the same as (29·1) except for the substitution of $l + \epsilon$ for l. It is not difficult to see that this equation can still be solved by the polynomial method of Sec. 29 and that it yields the eigenvalues

$$E_n(\epsilon) = -\frac{NhcZ^2}{(v + l + \epsilon + 1)^2} = \frac{E_n n^2}{(n + \epsilon)^2}. \quad (59\cdot9)$$

Hence

$$\left[\frac{dE_n(\epsilon)}{d\epsilon} \right]_{\epsilon=0} = -\frac{2E_n}{n} = \frac{2NhcZ^2}{n^3}. \quad (59\cdot10)$$

Elementary perturbation theory can also be used to evaluate $E_n(\epsilon)$ for small values of ϵ and thus $dE_n(\epsilon)/d\epsilon$ for the point $\epsilon = 0$. We obtain

$$\left[\frac{dE_n(\epsilon)}{d\epsilon} \right]_{\epsilon=0} = \left[\frac{dH_b}{d\epsilon} \right]_{\epsilon=0} = \frac{h^2}{8\pi^2\mu}(2l+1)\left(\overline{\frac{1}{r^2}} \right). \quad (59\cdot11)$$

Equating the right-hand members of Eqs. (59·10) and (59·11) gives

$$\overline{\frac{1}{r^2}} = 4\mu Z^2 \left(\frac{2\pi}{h} \right)^2 \frac{Nhc}{n^3(2l+1)}. \quad (59\cdot12)$$

Let $E_s(n,l,j)$ denote the mean value of U_s for the state whose quantum numbers are n, l, j. Let α denote the dimensionless fine-structure constant $\frac{2\pi c^2}{hc}\left(\cong \frac{1}{137} \right)$. Then Eqs. (58·6), (59·4), (59·7), and (59·12) yield

---

$$2E - \overline{rV'(r)} - \overline{2V(r)} = 0,$$

or, if T denotes the kinetic energy,

$$2\overline{T} = \overline{rV'(r)}.$$

This is the appropriate specialization of the *virial theorem*. For the special case of an hydrogenic atom this reduces to

$$Ze^2\left(\overline{\frac{1}{r}} \right) = -2E.$$

$$E_s\left(n, l, l + \frac{1}{2}\right) = \alpha^2 Z^4 \frac{Nhc}{n^3(l + 1)(2l + 1)}, \tag{59·13}$$

$$E_s\left(n, l, l - \frac{1}{2}\right) = -\alpha^2 Z^4 \frac{Nhc}{n^3 l(2l + 1)}. \tag{59·14}$$

The corresponding expression for the energy difference of two fine-structure levels belonging to the same parent n,l level is

$$\Delta E_s = \frac{\alpha^2 Z^4 Nhc}{n^3 l(l + 1)}. \tag{59·15}$$

The above formula with Z set equal to unity is directly applicable to the hydrogenic orbits of the alkalies and ions of the same structure, but in most cases the doublets for these states are too narrow to be measurable. In the case of almost hydrogenic states (slightly penetrating orbits) it can be used with the substitution of an appropriate effective value of the nuclear charge. In order to adapt the formula to the doublets of deeply penetrating orbits it is useful to employ a model in which the potential-energy function $V(r)$ follows a Coulomb law with an effective nuclear charge ze ($z = 1$ for neutral atoms) outside the core radius r_1 and the same law with a larger effective nuclear charge $Z_i e$ inside r_1. The average value of $V'(r)/r$ [cf. Eq. (58·5)] is then a suitable weighted average of the values ze^2/r^3 and $Z_i e^2/r^3$ appropriate to the outer and inner parts of the field.[1] Assuming that Z_i is much larger than z one obtains the Landé formula for the doublet spacing, viz.,

$$\Delta E_s = \frac{\alpha^2 Z_i^2 z^2 Nhc}{n^{*3} l(l + 1)}. \tag{59·16}$$

Here n^* denotes the effective quantum number defined by (56·8). This equation checks very satisfactorily with the experiment data,[1] and reflects clearly the previously mentioned rapid decrease in the doublet spacing with increasing values of n and l.

60. THE APPROXIMATE RELATIVISTIC THEORY OF THE HYDROGEN ATOM

Of course the hydrogen atom and the structurally similar ions He^+, Li^{++} form the simplest atoms with a single valence electron. The above theory should apply most accurately to them. In comparing the theory with the observed energy-level systems of such hydrogenic structures we must remember, however, that in first approximation the energy levels of hydrogen for a given total quantum number and different values of l are coincident. We should therefore expect the group of levels due to spin-orbit interaction for any given value of n to contain

---

[1] *Cf.* Pauling and Goudsmit, *The Structure of Line Spectra*, pp. 60–63, New York, 1930, for details.

$2n - 1$ components, *i.e.*, one for the s state and two for each possible value of l greater than zero. Actually the number is less than this due to the existence of a relativistic effect which we have so far ignored and to a residual degeneracy which eludes the combined action of the spin-orbit and relativistic terms in the Hamiltonian. The relativity correction in question is unimportant for the alkalies because it is small compared with the uncertainty in the theory due to the lack of an exact potential function $V(r)$ and to the essential inaccuracy of the single-electron model we have used. It is most important for the discussion of the low-lying X-ray energy levels.

The "pre-Dirac" relativistic equation for a single electron, (6·8), is not in the standard form $(H - E)\Psi = 0$ and is not easily put into it. It differs, however, from the non-relativistic equation (5·5) only in the addition of a term $\left(\dfrac{2\pi}{h}\right)^2 (E - V)^2 \dfrac{\Psi}{\mu c^2}$ which is ordinarily small compared with the other terms. Strictly speaking we cannot add this term to the ordinary Hamiltonian because it contains the unknown parameter E. We know the approximate value of E, however, by solving the equation without the $(E - V)^2$ term and can correct this by substituting the approximate value, say E_0, for E in that term. We thus obtain the equation[1]

$$\left\{H_0 - \frac{[E_0 - V(r)]^2}{2\mu c^2} - E\right\}\Psi = 0 \qquad (60\cdot1)$$

in which $H_0 - \dfrac{[E_0 - V(r)]^2}{2\mu c^2}$ plays the part of the perturbed Hamiltonian.

The relativistic energy-level correction E_R for a state with total quantum number n and azimuthal quantum number l is then the mean value of $-\dfrac{(E_0 - V)^2}{2\mu c^2}$ for the state in question. In the case of a hydrogenic atom with nuclear charge Ze,

$$E_R = -\frac{1}{2\mu c^2}\overline{\left(E_0 + \frac{Ze^2}{r}\right)^2}$$

$$= -\frac{1}{2\mu c^2}\left(E_0{}^2 + 2E_0 Ze^2\overline{\frac{1}{r}} + Z^2 e^4\overline{\frac{1}{r^2}}\right).$$

The mean values of $1/r^2$ and $1/r$ are given in Eq. (59·12) and footnote 4, p. 505, respectively. Hence

$$E_R = -\frac{\alpha^2 Z^4 Nhc}{n^3}\left[\frac{1}{l + \frac{1}{2}} - \frac{3}{4n}\right]. \qquad (60\cdot2)$$

[1] The eigenfunctions of this equation for different values of E_0 are not rigorously orthogonal.

Adding this to the spin correction we obtain the following approximate expression for the combined energy-level correction of relativity and the spin-orbit interaction:

$$E_{R+s}(n, l, l + \tfrac{1}{2}) = -\frac{\alpha^2 Z^4 N h c}{n^3}\left[\frac{1}{l+1} - \frac{3}{4n}\right], \tag{60·3}$$

$$E_{R+s}(n, l, l - \tfrac{1}{2}) = -\frac{\alpha^2 Z^4 N h c}{n^3}\left[\frac{1}{l} - \frac{3}{4n}\right]. \tag{60·4}$$

In this approximation, therefore, despite the perturbations introduced, there is a residual degeneracy due to the fact that

$$E(n, l, l + \tfrac{1}{2}) = E(n, l + 1, l + \tfrac{1}{2}). \tag{60·5}$$

As a result of this residual degeneracy it is appropriate to label the energy levels of hydrogen-like atoms by the corresponding j values rather than by the l values. We can then replace (60·3) and (60·4) by

$$E_{R+s}(n, j - \tfrac{1}{2}, j) = E_{R+s}(n, j + \tfrac{1}{2}, j) = -\frac{\alpha^2 Z^4 N h c}{n^3}\left[\frac{1}{j+\tfrac{1}{2}} - \frac{3}{4n}\right]. \tag{60·6}$$

The degeneracy is not removed by a more accurate theory as proved by Dirac's exact treatment of hydrogenic atoms on the basis of his rigorously relativistic quantum theory of the electron.[1]

The agreement between (60·6) and the observed spectra of H and He⁺ is too well known to require comment here.

The formulas for spin-orbit interaction and for the relativity correction are both of great importance in connection with X-ray energy levels. These are the imperfectly quantized energies of the ions formed by the removal of an inner electron from a relatively heavy atom. Their differences can be approximately identified with the negatives of the corresponding differences of the energy of the electron removed, if the latter is treated like the optical electron of the alkalies using a single-electron model in which the effective potential energy $V(r)$ is made up of the nuclear potential energy and an additional radial function to take the place of the average interaction potential with the other electrons. $V(r)$ is roughly hydrogenic for the inner electrons due to the dominance of the strong nuclear field. Hence Eqs. (59·14) and (60·2) can be used with appropriate effective values of Z. However, on account of the large values of Z_{eff} in the highest levels of the heavy atoms, our approximate treatment of the relativity correlation is not altogether satisfactory. The order of the fine-structure levels is, of course, reversed when we pass from the alkali spectra to the alkali-like X-ray spectra.

[1] P. A. M. DIRAC, *P.Q.M.*, 1st ed., Chap. XIII; 2d ed., Chap. XII.

61. THE PAULI WAVE-MECHANICAL FORMULATION OF THE THEORY OF ELECTRON SPIN

61a. Nature of the Configuration Space and Wave Functions.—Up to this point our application of quantum mechanics to the electron-spin hypothesis has been based on formal analogy for the purpose of arriving as quickly as possible at a quantitative check with the observed fine structure of the alkali spectra. Before applying the theory to the study of atoms with more than one valence electron we pause to develop a more complete quantum-mechanical formulation of the ideas roughly outlined in Sec. 58.

A very natural procedure for setting up a complete spin theory in wave-mechanical form would be to treat the electron as a rotating charged spherical rigid body whose position is fixed by Eulerian angles θ, φ, ψ. We should then use a wave equation for the free rotational motion of the electron like that of a symmetrical top with three equal moments of inertia. This procedure leads, however, to the same set of eigenvalues for the resultant angular momentum and for the component parallel to any fixed axis as we get for a system of particles. It does not give the half-integral spin quantum numbers demanded by the spectroscopic facts, nor would it give the single eigenvalue of the resultant S^2 which is observed experimentally. We might perhaps account for the absence of many empirical values of S^2 as a result of the impossibility of exciting higher internal-energy states, rather than as an evidence of their non-existence; but the half-integral quantum numbers impose a well-nigh fatal objection to any attempt to write out probability amplitudes with the Eulerian angles as arguments.[1] This fact is not very surprising when we reflect that the angles in question are essentially unobservable. We have, in fact, no use for such probability amplitudes except to complete the formal parallelism between the theory of spin angular momentum and that of orbital angular momentum.

Pauli has shown, however, that it is possible to give the electron-spin theory a formulation in terms of probability amplitudes without introducing coordinates which specify the orientation of the electron or that of its spin axis.[2] To this end we introduce the eigenvalues of spin angular-momentum components, or multiples of them, as new arguments for the wave functions. We have already become familiar in Sec. 36 with probability amplitudes some of whose arguments have discrete spectra. Only the eigenvalues of normally commuting dynamical variables can appear as simultaneous arguments in any probability amplitude. We attribute to the components of the spin the commutation properties we have already found to be characteristic of the orbital

[1] *Cf.*, however, C. G. DARWIN, *Proc. Roy. Soc.* **A115**, 1 (1927).

[2] W. PAULI, JR., *Zeits. f. Physik* **43**, 601 (1927).

angular momentum. Hence we can use at most the eigenvalues of one of the spin components in any one wave function. It is customary to choose S_z for this purpose. The square of the resultant spin angular momentum could be introduced as an additional argument, but this is unnecessary if we make the assumption that only one eigenvalue exists.

We assume that each of the three dynamical variables S_x, S_y, S_z whose operators we wish to define has just the two eigenvalues $\pm\frac{1}{2}\frac{h}{2\pi}$. It will be convenient to introduce two alternative symbols, *viz.*, σ and m_s, for the quantity $2\pi S_z'/h$, which we call the quantum number of S_z. Thus either σ or m_s denotes a quantity which can take on just two values, *viz.*, $\pm\frac{1}{2}$. In place of the eigenvalue S_z' itself we ordinarily use the quantum number σ as the fourth argument, or spin coordinate, of the probability amplitude of a single electron. Let $\alpha(\sigma)$, or $\delta_{\sigma,+\frac{1}{2}}$, denote that function of σ which is unity when σ is $+\frac{1}{2}$ and zero otherwise. Let $\beta(\sigma)$, or $\delta_{\sigma,-\frac{1}{2}}$, denote that function of σ which is unity when σ is $-\frac{1}{2}$ and zero otherwise. Any probability amplitude for the coordinates x, y, z, S_z can then be written in the form

$$\psi(x',y',z',\sigma) = u_\alpha(x',y',z')\alpha(\sigma) + u_\beta(x',y',z')\beta(\sigma). \qquad (61\cdot1)$$

Here u_α and u_β are simply two different physically admissible functions of the space coordinates x', y', z'. If we wish to indicate that a wave function is an eigenfunction of S_z with the eigenvalue S_z', we introduce m_s as the quantum number corresponding to S_z' and attach it as a subscript to the ψ symbol. Thus

$$\psi_{m_s}(x',y',z',\sigma) = u(x',y',z')\delta_{\sigma,m_s}$$

denotes such an eigenfunction.

In harmony with the procedures of Sec. 36 we assume the normalization condition

$$\sum_{x',y',z',\sigma} |\psi(x',y',z',\sigma)|^2 = 1,$$

or

$$\sum_{\sigma=-\frac{1}{2}}^{+\frac{1}{2}} \iiint \{|u_\alpha|^2\alpha(\sigma)^2 + |u_\beta|^2\beta(\sigma)^2$$
$$+ (u_\alpha u_\beta{}^* + u_\alpha{}^* u_\beta)\alpha(\sigma)\beta(\sigma)\}dx'dy'dz' = 1.$$

This reduces at once to

$$\iiint_\infty [|u_\alpha|^2 + |u_\beta|^2]dx'dy'dz' = 1. \qquad (61\cdot2)$$

With this normalization we shall interpret $|u_\alpha|^2 dx'dy'dz'$ as the probability that an electron lies in the volume element $dx'dy'dz'$ and that at the same

time σ has the value $\frac{1}{2}$. $|u_\beta|^2 dx' dy' dz'$ is the corresponding probability for the opposite value of σ.   Thus the introduction of the spin component S_z in effect replaces the three-argument wave function of the spinless electron by two such functions.

The reader will readily verify that if F is an operator function of the coordinates x, y, z and their momenta and has an adjoint $F^\dagger$ with respect to class D functions in x', y', z' space, it has the same adjoint with respect to class D and the enlarged x', y', z', σ space.   If F is Hermitian or unitary in x', y', z' space it retains these properties in x', y', z', σ space.

It is important to note at once that if we introduce the new wave functions into the old three-dimensional Schrödinger equation $H\psi = E\psi$ and ask for eigenvalues and eigenfunctions we obtain the same eigenvalues and multiples of the old eigenfunctions.   Thus it is necessary that

$$Hu_\alpha\alpha(\sigma) + Hu_\beta\beta(\sigma) = Eu_\alpha\alpha(\sigma) + Eu_\beta\beta(\sigma). \tag{61·3}$$

Multiplying this equation by $\alpha(\sigma)$ and summing up over the two possible values of σ we obtain $Hu_\alpha = Eu_\alpha$ with a similar equation for u_β.   Thus both u_α and u_β must be eigenfunctions if ψ is one.   Conversely, if u_α and u_β are eigenfunctions with the same eigenvalue, it follows that ψ is an eigenfunction.   If ψ is a simultaneous eigenfunction of all members of a complete set of commuting observables, say H, $\mathcal{L}^2$, $\mathcal{L}_z$, we can always write

$$\psi = u_{E,l,m_l}(x',y',z')[c_\alpha\alpha(\sigma) + c_\beta\beta(\sigma)], \tag{61·4}$$

where c_α and c_β are complex numbers.   If u_{E,l,m_l} is normalized in the usual way, c_α and c_β are subject to the normalization condition

$$|c_\alpha|^2 + |c_\beta|^2 = 1. \tag{61·5}$$

61b. Preliminary Discussion of Spin Operators and Spin Matrices.— If we set either c_α or c_β equal to zero, ψ_{E,l,m_l} becomes an eigenfunction of S_z.   Thus we obtain a complete set of simultaneous eigenfunctions of H, $\mathcal{L}^2$, $\mathcal{L}_z$, and S_z:

$$\psi_{E,l,m_l,m_s}(x',y',z',\sigma) = u_{E,l,m_l}(x',y',z')\delta_{\sigma,m_s}. \tag{61·6}$$

As an arbitrary physically admissible function of x', y', z', σ is expansible in terms of the above set, the equation

$$S_z\psi_{E,l,m_l,m_s} = m_s\left(\frac{h}{2\pi}\right)\psi_{E,l,m_l,m_s} \tag{61·7}$$

is sufficient to define S_z.   Clearly S_z is a dynamical variable which commutes with all operators which are functions of the Cartesian coordinates x', y', z' and their momenta, but are independent of the spin.   It has the property that, when applied to the product of a function of the space coordinates x', y', z' and a function of the spin coordinate σ, it acts only

on the latter. In other words we can treat x', y', z' as parameters in applying S_z.

We assume that the other spin operators, S_x and S_y share the same general properties as S_z, but do not commute with S_z. Since they act only on the spin functions, they are determined by four-element matrices. Thus

$$S_x \beta(\sigma) = \beta(\sigma) S_x(\beta,\beta) + \alpha(\sigma) S_x(\alpha,\beta), \left.\right\}$$
$$S_x \alpha(\sigma) = \beta(\sigma) S_x(\beta,\alpha) + \alpha(\sigma) S_x(\alpha,\alpha). \left.\right\} \tag{61.8}$$

In a matrix scheme in which S_z and a set of spin-free operators $\gamma_1, \gamma_2, \cdots$ are diagonal, the matrices of all three spin operators are not only diagonal with respect to the γ's but have elements entirely independent of the values of the γ's. Thus, if E' and E'' are discrete energy levels of a spin-free Hamiltonian,

$$(E', l'm_l', \tfrac{1}{2} |S_x| E'', l'', m_l'', -\tfrac{1}{2}) = \sum_{x',y',z',\sigma} u^*_{E',l',m_{l'}}(x',y',z') \alpha(\sigma)$$

$$\times S_x u_{E'',l'',m_{l''}}(x',y',z')\beta(\sigma) = \delta_{E',E''}\delta_{l',l''}\delta_{m_{l'},m_{l''}} \sum_\sigma \alpha(\sigma) S_x \beta(\sigma)$$

$$= \delta_{E',E''}\delta_{l',l''}\delta_{m_{l'},m_{l''}} S_x(\alpha,\beta). \tag{61.9}$$

We can accordingly suppress the indices E, l, m_l and treat S_x, S_y, and S_z as (2×2)-element matrices whose rows and columns are labeled by m_s values.

The operators S_x, S_y are determined by their matrices. In order to fix the latter it is convenient to make use of an analysis which is equally applicable to the matrices of the orbital angular momentum, the spin angular momentum, and the resultant $\overrightarrow{g}$ of the spin and orbital angular momenta. Hence it is worth while to treat the general case first and to consider the specialization appropriate to the spin later.[1]

We assume (*cf.* Sec. 58c) that the rules for the commutation of S_x, S_y, S_z among themselves are the same as for $\mathcal{L}_x$, $\mathcal{L}_y$, $\mathcal{L}_z$. They then form a special case of the operators α, β, γ of Sec. 40f. The eigenvalues we have postulated for the components of $\overrightarrow{s}$ and for S^2 are in harmony with the results obtained in Sec. 40f for the eigenvalues of γ and ω^2. We accordingly proceed to the general problem of working out the matrices of the operators α, β, γ, ω^2.

Let ρ denote the dynamical variable $\alpha + i\beta$, whose analogue $\mathcal{L}_x + i\mathcal{L}_y$ proved so useful in Sec. 40f. Due to the Hermitian character of α and β

[1] M. BORN, W. HEISENBERG, and P. JORDAN, *Zeits. f. Physik* **35**, 557 (1926); P. A. M. DIRAC, *P.Q.M.*, 2d ed. §§48, 49.

the dynamical variable $\alpha - i\beta$ is adjoint to ρ and can be labeled $\rho^\dagger$. It follows from (40·16) that

$$\gamma\varrho = \varrho\left(\gamma + \frac{h}{2\pi}\mathbf{I}\right),$$
(61·10)

where $\mathbf{I}$ denotes the unit matrix.

We assume a matrix scheme based on simultaneous eigenfunctions of γ, ω^2 and such additional commuting Hermitian operators μ_1, μ_2, $\cdots$ as are needed to make a complete set of normally commuting variables. Let l and m denote the quantum numbers of ω^2 and γ as defined on p. 316. It is assumed that β and γ as well as α commute with the μ's. The matrices $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\varrho}$, $\boldsymbol{\varrho}^\dagger$, $\boldsymbol{\omega}^2$ will then be diagonal with respect to l (cf. Sec. 43, p. 351) and the μ's. Introducing an integer τ, each value of which denotes a set of simultaneous eigenvalues of the μ's we indicate a typical element of ϱ which is diagonal with respect to l and the μ's by $\rho_{\tau l}(m',m'')$. All elements of ϱ, $\varrho^\dagger$, α, β, γ, ω^2 not expressible in this form are zero. The totality of the elements of ϱ for a single pair of values of τ and l form a square matrix which is conveniently designated by $\varrho_{\tau l}$. Equation (61·10) now leads to

$$\gamma_{\tau l}\varrho_{\tau l} = \varrho_{\tau l}\left(\gamma_{\tau l} + \frac{h}{2\pi}\mathbf{I}_{\tau l}\right).$$
(61·11)

Here $\mathbf{I}_{\tau l}$ is a unit matrix with the same number of rows and columns as $\varrho_{\tau l}$, $\gamma_{\tau l}$, etc. As γ has only diagonal elements of value $mh/2\pi$, this matrix equation is equivalent to

$$(m' - m'' - 1)\rho_{\tau l}(m',m'') = 0.$$
(61·12)

Thus all elements of $\varrho_{\tau l}$ vanish except those on that line parallel to the principal diagonal for which $m' = m'' + 1$. Let

$$\rho_{\tau l}(m + 1,m) = a_{\tau l}(m).$$

Then $\rho_{\tau l}^\dagger(m,m + 1) = a_{\tau l}(m)^*$. All other elements vanish.

In order to evaluate $a_{\tau l}(m)$ we make use of the matrix equation

$$\varrho_{\tau l}^\dagger\varrho_{\tau l} = \omega_{\tau l}^2 - \gamma_{\tau l}^2 - \gamma_{\tau l}\frac{h}{2\pi},$$
(61·13)

which is a corollary on (40·18). Introducing appropriate element values into (61·13) we obtain

$$|a_{\tau l}(m)|^2 = \left(\frac{h}{2\pi}\right)^2 [l(l + 1) - m(m + 1)].$$
(61·14)

It follows that

$$a_{\tau l}(m) \equiv (\tau,l,m + 1|\rho|\tau,l,m) = e^{i\nu(\tau,l,m)}\frac{h}{2\pi}\sqrt{l(l + 1) - m(m + 1)},$$
(61·15)

where $\nu(\tau,l,m)$ is a phase constant which cannot be determined by matrix

algebra.   It has definite values for every set of indices when the matrix scheme is related to a set of wave functions $\psi_{\tau l m}$ with definite phases according to the basic formula (44·2).   As the phases of the wave functions are arbitrary, the constants $\nu(\tau,l,m)$ are also arbitrary in a sense and are usually set equal to zero.

With the above choice of phases the relations $\alpha = \frac{1}{2}(\rho + \rho^\dagger)$ and $\beta = \frac{1}{2}i(-\rho + \rho^\dagger)$ yield

$$\alpha_{\tau l}(m',m'') = \frac{h}{4\pi}\{\sqrt{l(l+1) - m''(m''+1)}\,\delta_{m',m''+1}$$
$$+ \sqrt{l(l+1) - m'(m'+1)}\,\delta_{m'+1,m''}\}, \quad (61\cdot16)$$

$$\beta_{\tau l}(m',m'') = i\frac{h}{4\pi}\{-\sqrt{l(l+1) - m''(m''+1)}\,\delta_{m',m''+1}$$
$$+ \sqrt{l(l+1) - m'(m'+1)}\,\delta_{m'+1,m''}\}. \quad (61\cdot17)$$

Let us now apply these general results to the spin angular momentum of an individual electron.   In this case we replace the symbol m by m_s and restrict its values to $+\frac{1}{2}$ and $-\frac{1}{2}$.   We replace the symbol l by s and give it the single value $+\frac{1}{2}$ in accordance with the assumption of Sec. 58d, p. 500 and the analysis of Sec. 40f.   We can suppress s as an index since it has only the one value.   The matrices $S_{x\tau} + iS_{y\tau}$ and $S_{x\tau} - iS_{y\tau}$ have only one non-vanishing element.   Thus

$$(m_s'|S_{x\tau} + iS_{y\tau}|m_s'') = \frac{h}{2\pi}\sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right) - m_s''(m_s''+1)}\,\delta_{m_s',m_s''+1}$$

$$= \frac{h}{2\pi}\delta_{\frac{1}{2},m_s'}\delta_{-\frac{1}{2},m_s''}.$$

Let us arrange the m_s values for the different rows and columns in order of increasing magnitude as follows:[1]

---

[1] In comparing Eqs. (61·18) and (61·19) with the corresponding equations given by other authors the reader must exercise circumspection.   Thus Pauli in his original article (*loc. cit.*, p. 510) and his article in Geiger and Scheel's *Handbuch der Physik*, **24**, Part I, reverses the ordering of rows and columns used here.   Wigner, *Gruppentheorie und ihre Anwendung*, p. 251, Berlin, 1931, has interchanged the x and y components of the matrix of S.

Then,

$$S_{x\tau} + iS_{y\tau} = \frac{h}{2\pi}\begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix}, \qquad S_{x\tau} - iS_{y\tau} = \frac{h}{2\pi}\begin{Vmatrix} 0 & 1 \\ 0 & 0 \end{Vmatrix}. \qquad (61\cdot18)$$

$$S_{x\tau} = \frac{h}{4\pi}\begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}, \qquad S_{y\tau} = \frac{h}{4\pi}\begin{Vmatrix} 0 & i \\ -i & 0 \end{Vmatrix}, \qquad S_{z\tau} = \frac{h}{4\pi}\begin{Vmatrix} -1 & 0 \\ 0 & 1 \end{Vmatrix}. \qquad (61\cdot19)$$

Having determined the spin matrices, we return to the corresponding operators. The matrix component $S_{x\tau}(\alpha,\beta)$ of Eq. (61·9) is the component of $S_{x\tau}$ for which the row index m_s' is $+\frac{1}{2}$ and the column index m_s'' is $-\frac{1}{2}$. Thus the arrangement of the elements in Eqs. (61·8) is transposed with respect to the arrangements of Eqs. (61·18) and (61·19). With this understanding, (61·8) is the type form for the conversion of all the spin matrices into operators.

We have assumed hitherto that the spin operators act only on the spin functions when a probability amplitude is factored into the product of a spin factor and a position factor. It is formally possible, however, to get the same results with a different point of view. Consider the application of the operator S_x to the general wave function of (61·1). By means of (61·8) we obtain

$$\begin{aligned} S_x\psi &= S_x[u_\beta\beta(\sigma) + u_\alpha\alpha(\sigma)] \\ &= u_\beta[\beta S_x(\beta,\beta) + \alpha S_x(\alpha,\beta)] + u_\alpha[\beta S_x(\beta,\alpha) + \alpha S_x(\alpha,\alpha)] \\ &= \beta[S_x(\beta,\beta)u_\beta + S_x(\beta,\alpha)u_\alpha] + \alpha[S_x(\alpha,\beta)u_\beta + S_x(\alpha,\alpha)u_\alpha]. \qquad (61\cdot20) \end{aligned}$$

Thus S_x can be thought of as transforming u_β into $S_x(\beta,\beta)u_\beta + S_x(\beta,\alpha)u_\alpha$ and u_α into $S_x(\alpha,\beta)u_\beta + S_x(\alpha,\alpha)u_\alpha$. Inserting the actual values of these matrix elements, we see that from our present point of view S_x transforms u_β into $hu_\alpha/4\pi$ and u_α into $hu_\beta/4\pi$. Similarly S_y transforms u_β into $\dfrac{ihu_\alpha}{4\pi}$ and u_α into $-\dfrac{ihu}{4\pi}\beta$. Finally S_z transforms u_β into $-\dfrac{hu_\beta}{4\pi}$ and u_α into $hu_\alpha/4\pi$. This interpretation of the operators is very useful in practice.

If we like, we can regard the two functions $u_\beta(x',y',z')$ and $u_\alpha(x',y',z')$ as the elements of a (2×1)-element matrix function of the coordinates

$$\psi(x',y',z') = \begin{Vmatrix} u_\beta \\ u_\alpha \end{Vmatrix}. \qquad (61\cdot21)$$

The above rules then say that

$$S_x\psi = \begin{Vmatrix} S_x(\beta,\beta)u_\beta + S_x(\beta,\alpha)u_\alpha \\ S_x(\alpha,\beta)u_\beta + S_x(\alpha,\alpha)u_\alpha \end{Vmatrix} = S_{x\tau}\psi. \qquad (61\cdot22)$$

(4×1)-component matrix wave functions similar to the above are used in Dirac's relativistic quantum theory of the electron.

It is of considerable importance to be able to make transformations of the spin wave functions and spin operators from one set of coordinate

axes to another. The theory of these transformations was first worked out by Pauli.[1]

Let the new coordinate axes $\bar{x}$, $\bar{y}$, $\bar{z}$ be derived from the original set x, y, z by rotation through the Eulerian angles φ, θ, ψ (*cf.* Fig. 14), Sec. 34*f*. Since the angular-momentum operators $\mathcal{L}_x$, $\mathcal{L}_y$, $\mathcal{L}_z$, which we have used as models for the construction of S_x, S_y, S_z, form a vector—as indicated by the notation $\overrightarrow{\mathcal{L}}$—we postulate that $\overrightarrow{S}$ is a vector, *i.e.*, that

$$\begin{aligned}
S_{\bar{x}} &= S_x \cos(x,\bar{x}) + S_y \cos(y,\bar{x}) + S_z \cos(z,\bar{x}), \\
S_{\bar{y}} &= S_x \cos(x,\bar{y}) + S_y \cos(y,\bar{y}) + S_z \cos(z,\bar{y}), \\
S_{\bar{z}} &= S_x \cos(x,\bar{z}) + S_y \cos(y,\bar{z}) + S_z \cos(z,\bar{z}).
\end{aligned} \qquad (61\cdot23)$$

Let T denote the operator which transforms an x,y,z spin wave function $(\sigma|)$ into the corresponding $\bar{x},\bar{y},\bar{z}$ spin wave function $(\bar{\sigma}|)$. The possible values of the argument $\bar{\sigma}$ are the eigenvalues of $S_{\bar{z}}$, each multiplied by $2\pi/h$. Of course these values must be $\pm\frac{1}{2}$ if one set of axes is to be fully equivalent to the other. Let $(\sigma|\bar{\sigma})$ denote the eigenfunction of $S_{\bar{z}}$ in the x,y,z system for the eigenvalue $\bar{\sigma}h/2\pi$. By the fundamental equation (36·76),

$$(\bar{\sigma}|) = \sum_{m_s} (\sigma|\bar{\sigma})^*(\sigma|) = \sum_{m_s} (\bar{\sigma}|\sigma)(\sigma|) \equiv T(\sigma|). \qquad (61\cdot24)$$

It follows from Eq. (36·5) that the transformed expressions for the operators $S_{\bar{x}}$, $S_{\bar{y}}$, $S_{\bar{z}}$, to be used in connection with the wave functions $(\bar{\sigma}|)$, are

$$\bar{S}_{\bar{x}} = T S_{\bar{x}} T^{-1}, \qquad \bar{S}_{\bar{y}} = T S_{\bar{y}} T^{-1}, \qquad \bar{S}_{\bar{z}} = T S_{\bar{z}} T^{-1}. \qquad (61\cdot25)$$

In order that the two sets of axes shall be fully equivalent it is necessary that $\bar{S}_{\bar{x}}$, $\bar{S}_{\bar{y}}$, $\bar{S}_{\bar{z}}$ shall be identical in form with S_x, S_y, S_z, respectively. The fact that this equivalence can be attained by a proper choice of the phases of the eigenfunctions $(\sigma|\bar{\sigma})$ is proof that despite the asymmetry of its external form the Pauli spin theory is in harmony with the isotropic character of space.[2]

Since the spin operators have purely discrete spectra with two eigenvalues each, the functions $(\sigma|\bar{\sigma})^*$, *i.e.*, $(\bar{\sigma}|\sigma)$, form a matrix with two rows and two columns which we call $\mathbf{T}$, thereby establishing a parallelism between (61·24) and the first of Eqs. (44·15). In place of Eqs. (61·25) we can use

$$\bar{\mathbf{S}}_{\bar{x}} = \mathbf{T} \mathbf{S}_{\bar{x}} \mathbf{T}^{-1} = \mathbf{T} \mathbf{S}_{\bar{x}} \mathbf{T}^\dagger, \qquad (61\cdot26)$$

---

[1] *Loc. cit.*, footnote 2, p. 510.

[2] A more complete discussion of this question with a derivation of the transformation matrix (61·30) is to be found in E. Wigner's *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*, Berlin, 1931.

with corresponding equations for $\bar{S}_{\bar{y}}$ and $\bar{S}_{\bar{z}}$. We have now to choose the unitary matrix $\mathbf{T}$ so that

$$\mathbf{T}S_{\bar{z}}\mathbf{T}^{-1} = S_x, \qquad \mathbf{T}S_{\bar{y}}\mathbf{T}^{-1} = S_y, \qquad \mathbf{T}S_{\bar{z}}\mathbf{T}^{-1} = S_z. \qquad (61\cdot27)$$

The existence of a solution of Eqs. (61·27) insures the existence of the eigenfunctions $(\sigma|\bar{\sigma})^*$ and defines the operator T of Eqs. (61·24) and (61·25).

It will suffice here to deal explicitly with the simple case where the $\bar{x},\bar{y},\bar{z}$ system is obtained from the x,y,z system by a rotation through an angle φ about the z axis. The $\bar{z}$ axis is then coincident with the z axis and the operators $S_{\bar{z}}$ and S_z are identical. It is nevertheless necessary to introduce a phase difference between the eigenfunctions of $S_{\bar{z}}$ and S_z in order to make the matrices $\bar{S}_{\bar{z}}$ and $\bar{S}_{\bar{y}}$ identical with S_x and S_y, respectively. In this case

$$S_{\bar{x}} = S_x \cos\varphi + S_y \sin\varphi = \frac{h}{4\pi}\left\|\begin{matrix} 0 & e^{i\varphi} \\ e^{-i\varphi} & 0 \end{matrix}\right\|, \qquad (61\cdot28)$$

$$S_{\bar{y}} = -S_x \sin\varphi + S_y \cos\varphi = \frac{h}{4\pi}\left\|\begin{matrix} 0 & e^{i(\varphi+\frac{\pi}{2})} \\ e^{-i(\varphi+\frac{\pi}{2})} & 0 \end{matrix}\right\|. \qquad (61\cdot29)$$

Since $S_{\bar{z}}$ is identical with the diagonal matrix S_z, the last of Eqs. (61·27) requires that $\mathbf{T}$ shall be a diagonal unitary matrix, *i.e.*, that it shall have the form

$$\mathbf{T} = \left\|\begin{matrix} e^{i\alpha} & 0 \\ 0 & e^{i\beta} \end{matrix}\right\|.$$

The remaining Eqs. (61·27) yield

$$T(\lambda,\lambda)S_{\bar{z}}(\lambda,\mu) = S_x(\lambda,\mu)T(\mu,\mu),$$
$$T(\lambda,\lambda)S_{\bar{y}}(\lambda,\mu) = S_y(\lambda,\mu)T(\mu,\mu).$$

Inserting the values of the matrix elements, we obtain the four equations

$$e^{i\alpha}e^{i\varphi} = e^{i\beta}, \qquad e^{i\beta}e^{-i\varphi} = e^{i\alpha},$$

$$e^{i\alpha}e^{i\left(\varphi+\frac{\pi}{2}\right)} = e^{i\left(\beta+\frac{\pi}{2}\right)}, \qquad e^{i\beta}e^{-i\left(\varphi+\frac{\pi}{2}\right)} = e^{i\left(\alpha-\frac{\pi}{2}\right)},$$

all of which are satisfied if we set $\beta - \alpha = \varphi$. Symmetry suggests that we give β and α the values $\pm\frac{\varphi}{2}$, respectively. From the definition of $\mathbf{T}$ it follows that the functions $(\sigma|\bar{\sigma})$ are

$$(\sigma|-\frac{1}{2}) = e^{\frac{i\varphi}{2}}\beta(\sigma), \qquad (\sigma|+\frac{1}{2}) = e^{-\frac{i\varphi}{2}}\alpha(\sigma).$$

The complete transformation matrix $\mathbf{T}(\varphi,\theta,\psi)$ for a rotation with the arbitrary Eulerian angles φ, θ, ψ, as given by Pauli[1] and Wigner,[2] is

[1] *Loc. cit.*, footnote 2, p. 510.
[2] *Loc. cit.*, footnote 1, p. 515.

$$\mathbf{T}(\varphi,\theta,\psi) \;=\; \left\| \begin{array}{cc} e^{-\frac{i\varphi}{2}} \cos \dfrac{1}{2}\theta\, e^{-\frac{i\psi}{2}} & -e^{-\frac{i\varphi}{2}} \sin \dfrac{\theta}{2}\, e^{\frac{i\psi}{2}} \\[2mm] e^{\frac{i\varphi}{2}} \sin \dfrac{1}{2}\theta\, e^{-\frac{i\psi}{2}} & e^{\frac{i\varphi}{2}} \cos \dfrac{1}{2}\theta\, e^{\frac{i\psi}{2}} \end{array} \right\| . \tag{61·30}$$

61c. Application of the Pauli Theory to the Alkali Doublets.—As a simple exercise let us reexamine the problem of the spin-orbit interaction for the alkali atoms in terms of the Pauli wave-mechanical theory. Let H_0 denote the unperturbed Hamiltonian of the atomic model used in Sec. 56, *i.e.*,

$$H_0 \;=\; -\frac{h^2}{8\pi^2\mu}\nabla^2 + V(r). \tag{61·31}$$

The complete Hamiltonian is then

$$H \;=\; H_0 + \frac{1}{2\mu^2 c^2}\left(\frac{1}{r}\frac{dV}{dr}\right)(\vec{\mathbf{s}}\cdot\vec{\mathcal{L}}) \tag{61·32}$$

if the relativistic corrections are omitted. We write down the corresponding Schrödinger equation, set $\psi = \alpha(\sigma)u_\alpha + \beta(\sigma)u_\beta$ and make use of the properties of the spin operators as formulated in Sec. 61*b* to obtain

$$\alpha(\sigma)\left\{(H_0 - E)u_\alpha + \left(\frac{1}{r}\frac{dV}{dr}\right)\frac{(\mathcal{L}_x u_\beta - i\mathcal{L}_y u_\beta + \mathcal{L}_z u_\alpha)h}{8\pi\mu^2 c^2}\right\}$$
$$+ \beta(\sigma)\left\{(H_0 - E)u_\beta + \left(\frac{1}{r}\frac{dV}{dr}\right)\frac{(\mathcal{L}_x u_\alpha + i\mathcal{L}_y u_\alpha - \mathcal{L}_z u_\beta)h}{8\pi\mu^2 c^2}\right\} = 0. \tag{61·33}$$

Since this equation is to hold for all values of σ, the cofactors of $\alpha(\sigma)$ and $\beta(\sigma)$ must vanish separately. Hence (61·33) is equivalent to the pair of simultaneous equations

$$\left.\begin{array}{l} (H_0 - E)u_\alpha + \left(\dfrac{1}{r}\dfrac{dV}{dr}\right)\dfrac{[(\mathcal{L}_x - i\mathcal{L}_y)u_\beta + \mathcal{L}_z u_\alpha]h}{8\pi^2\mu^2 c^2} = 0, \\[4mm] (H_0 - E)u_\beta + \left(\dfrac{1}{r}\dfrac{dV}{dr}\right)\dfrac{[(\mathcal{L}_x + i\mathcal{L}_y)u_\alpha - \mathcal{L}_z u_\beta]h}{8\pi^2\mu^2 c^2} = 0. \end{array}\right\} \tag{61·34}$$

To get the appropriate zero-order wave functions and the first-order energy corrections we can enter the equations with expansions such as

$$u_\alpha \;=\; \sum_{m_l=-l}^{+l} C^\alpha_{m_l} R_{nl}\Theta_{lm_l}\Phi_{m_l}, \tag{61·35}$$

$$u_\beta \;=\; \sum_{m_l=-l}^{+l} C^\beta_{m_l} R_{nl}\Theta_{lm_l}\Phi_{m_l}, \tag{61·36}$$

and attempt to determine appropriate values of the constants $C^\alpha_{m_l}$, $C^\beta_{m_l}$ by conventional perturbation theory. Or, we can seek first to find such

values of the coefficients as will make ψ from the beginning a simultaneous eigenfunction of H_0, $\mathcal{L}^2$ and the operators $\mathcal{J}^2$ and $\mathcal{J}_z$ defined by the equations

$$\mathcal{J}_x = \mathcal{L}_x + S_x, \quad \mathcal{J}_y = \mathcal{L}_y + S_y, \quad \mathcal{J}_z = \mathcal{L}_z + S_z, \quad \mathcal{J}^2 = \mathcal{J}_x{}^2 + \mathcal{J}_y{}^2 + \mathcal{J}_z{}^2 \tag{61·37}$$

[cf. Eq. (59·1)]. Let us choose the latter method.

Our first step must be to prove the usefulness of $\mathcal{J}_z$ and $\mathcal{J}^2$ by verifying the assumption that they commute with the perturbed Hamiltonian H of Eq. (61·32). Since H is symmetric with respect to the three coordinate axes, the commutivity of $\mathcal{J}_z$ with H implies the commutivity of $\mathcal{J}_x$, $\mathcal{J}_y$, $\mathcal{J}^2$ with H. Hence we need only to examine $\mathcal{J}_z$. We know that S_z commutes with $\mathcal{L}_z$ and with any function of the coordinates x, y, z. $\mathcal{L}_z$ commutes with any function of the radius [cf. Eq. (38·11)]. Hence

$$\frac{1}{r}\frac{dV}{dr}(\vec{S}\cdot\vec{\mathcal{L}})\mathcal{J}_z - \mathcal{J}_z\frac{1}{r}\frac{dV}{dr}(\vec{S}\cdot\vec{\mathcal{L}}) = \frac{1}{r}\frac{dV}{dr}[(\vec{S}\cdot\vec{\mathcal{L}})\mathcal{J}_z - \mathcal{J}_z(\vec{S}\cdot\vec{\mathcal{L}})]$$

$$= \frac{1}{r}\frac{dV}{dr}\{S_x\mathcal{L}_x(\mathcal{L}_z + S_z) - (\mathcal{L}_z + S_z)S_x\mathcal{L}_x + S_y\mathcal{L}_y(\mathcal{L}_z + S_z) - (\mathcal{L}_z + S_z)S_y\mathcal{L}_y\}$$

$$= \frac{1}{r}\frac{dV}{dr}\{S_x(\mathcal{L}_x\mathcal{L}_z - \mathcal{L}_z\mathcal{L}_x) - \mathcal{L}_y(S_zS_y - S_yS_z)$$

$$+ \mathcal{L}_x(S_xS_z - S_zS_x) - S_y(\mathcal{L}_z\mathcal{L}_y - \mathcal{L}_y\mathcal{L}_z)\} = 0. \tag{61·38}$$

This proves that $\mathcal{J}_z$ commutes with the perturbing terms in (61·32). Consequently $\mathcal{J}_z$, $\mathcal{J}_x$, $\mathcal{J}_y$, $\mathcal{J}^2$ commute with that term and with the complete Hamiltonian H.

The functions u_α, u_β of Eqs. (61·35) and (61·36) are from the beginning eigenfunctions of H_0 and $\mathcal{L}^2$. Consider next the operator $\mathcal{J}_z$. Every simultaneous eigenfunction of $\mathcal{L}_z$ and S_z is also an eigenfunction of $\mathcal{J}_z$. Let Y_{l,m_l} denote the product $\Theta_{lm_l}\Phi_{m_l}$. We can then say that $R_{nl}Y_{lm_l}\alpha(\sigma)$ is an eigenfunction of $\mathcal{J}_z$ with the eigenvalue

$$\mathcal{J}_z' = \frac{mh}{2\pi} \equiv \frac{(m_l + \frac{1}{2})h}{2\pi}.$$

The most general simultaneous eigenfunction of H_0, $\mathcal{L}^2$, and $\mathcal{J}_z$ is

$$\psi_{n,l,m} = R_{nl}[C_m^\alpha \alpha(\sigma)Y_{l,m-\frac{1}{2}} + C_m^\beta \beta(\sigma)Y_{l,m+\frac{1}{2}}]. \tag{61·39}$$

Let us next try to choose C_m^α and C_m^β so that the above expression will become an eigenfunction of $\mathcal{J}^2$. Using j as the quantum number of $\mathcal{J}^2$, we denote the resulting simultaneous eigenfunction by $\psi_{n,l,j,m}$. In view of (59·2)

$$\frac{4\pi^2}{h^2}\mathcal{J}^2\psi_{n,l,j,m} = [l(l+1) + s(s+1)]\psi_{n,l,j,m}$$

$$+ \frac{2\pi}{h}\{\alpha(\sigma)[C_m^\beta(\mathcal{L}_x - i\mathcal{L}_y)Y_{l,m+\frac{1}{2}} + C_m^\alpha\mathcal{L}_zY_{l,m-\frac{1}{2}}]$$

$$+ \beta(\sigma)[C_m^\alpha(\mathcal{L}_x + i\mathcal{L}_y)Y_{l,m-\frac{1}{2}} - C_m^\beta\mathcal{L}_zY_{l,m+\frac{1}{2}}]\}. \tag{61·40}$$

The matrices $L_x + iL_y$ and $L_x - iL_y$ are special cases of the matrices ϱ and $\varrho^\dagger$ of p. 514. The only nonvanishing elements of the former are

$$(m_l + 1|\mathcal{L}_x + i\mathcal{L}_y|m_l) = \left(\frac{h}{2\pi}\right)\left[l(l+1) - \left(m - \frac{1}{2}\right)\left(m + \frac{1}{2}\right)\right]^{\frac{1}{2}}$$

[cf. Eq. (61·16)].   Hence

$$(\mathcal{L}_x + i\mathcal{L}_y)Y_{l,m-\frac{1}{2}} = \left(\frac{h}{2\pi}\right)\left[l(l+1) - m^2 + \frac{1}{4}\right]^{\frac{1}{2}}Y_{l,m+\frac{1}{2}}.$$

Similarly,

$$(\mathcal{L}_x - i\mathcal{L}_y)Y_{l,m+\frac{1}{2}} = \left(\frac{h}{2\pi}\right)\left[l(l+1) - m^2 + \frac{1}{4}\right]^{\frac{1}{2}}Y_{l,m-\frac{1}{2}}.$$

We have to solve

$$\left[\frac{4\pi^2}{h^2}\mathcal{J}^2 - j(j+1)\right]\psi_{n,l,j,m} = 0.$$

Setting the cofactors of $\alpha(\sigma)$ and $\beta(\sigma)$ equal to zero, and introducing the contractions

$$Q \equiv l(l+1) + \tfrac{3}{4} - j(j+1), \qquad Z \equiv [l(l+1) - m^2 + \tfrac{1}{4}]^{\frac{1}{2}}, \quad (61\cdot41)$$

we derive

$$\left.\begin{array}{l}(Q + m - \tfrac{1}{2})C_m^\alpha + ZC_m^\beta = 0, \\ ZC_m^\alpha + (Q - m - \tfrac{1}{2})C_m^\beta = 0.\end{array}\right\} \qquad (61\cdot42)$$

The condition for a nontrivial solution is

$$(Q - \tfrac{1}{2} + m)(Q - \tfrac{1}{2} - m) - Z^2 = 0.$$

Hence Q must have one of the two values $l+1$ or $-l$.   In the former case $j = l - \tfrac{1}{2}$ and

$$\frac{C_m^\alpha}{C_m^\beta} = -\left(\frac{l + \tfrac{1}{2} - m}{l + \tfrac{1}{2} + m}\right)^{\frac{1}{2}}. \qquad (61\cdot43)$$

In the latter case $j = l + \tfrac{1}{2}$ and

$$\frac{C_m^\alpha}{C_m^\beta} = \left(\frac{l + \tfrac{1}{2} + m}{l + \tfrac{1}{2} - m}\right)^{\frac{1}{2}}. \qquad (61\cdot44)$$

It follows from the normalization condition (61·5) that if $j = l - \tfrac{1}{2}$ and we give C_m^α a positive real value

$$C_m^\alpha = \left(\frac{l + \tfrac{1}{2} - m}{2l + 1}\right)^{\frac{1}{2}}, \qquad C_m^\beta = -\left(\frac{l + \tfrac{1}{2} + m}{2l + 1}\right)^{\frac{1}{2}},$$

while, if $j = l + \tfrac{1}{2}$,

$$C_m^\alpha = \left(\frac{l + \tfrac{1}{2} + m}{2l + 1}\right)^{\frac{1}{2}}, \qquad C_m^\beta = \left(\frac{l + \tfrac{1}{2} - m}{2l + 1}\right)^{\frac{1}{2}}.$$

The resulting zero-order wave functions of H are

$$
\left.
\begin{aligned}
\psi_{n,l,l-\frac{1}{2},m} &= \frac{R_{nl}}{(2l+1)^{\frac{1}{2}}} \left\{ \left(l + \frac{1}{2} - m \right)^{\frac{1}{2}} Y_{l,m-\frac{1}{2}} \alpha(\sigma) \right. \\
&\qquad\qquad \left. - \left(l + \frac{1}{2} + m \right)^{\frac{1}{2}} Y_{l,m+\frac{1}{2}} \beta(\sigma) \right\}, \\
\psi_{n,l,l+\frac{1}{2},m} &= \frac{R_{nl}}{(2l+1)^{\frac{1}{2}}} \left\{ \left(l + \frac{1}{2} + m \right)^{\frac{1}{2}} Y_{l,m-\frac{1}{2}} \alpha(\sigma) \right. \\
&\qquad\qquad \left. + \left(l + \frac{1}{2} - m \right)^{\frac{1}{2}} Y_{l,m+\frac{1}{2}} \beta(\sigma) \right\}.
\end{aligned}
\right\}
\quad (61 \cdot 45)
$$

Let us now apply the perturbation theory of Sec. 48a to determine the first-order energy corrections. As $\mathcal{L}^2$ and $\mathcal{L}_z$ commute with the perturbing term of the Hamiltonian (61·32) the secular determinant of (48·7) has no nonvanishing nondiagonal elements. The energy corrections are the diagonal elements of the matrix of $\left(\dfrac{1}{r} \right) \dfrac{dV}{dr} \dfrac{(\vec{s} \cdot \vec{\mathcal{L}})}{2\mu^2 c^2}$. Thus we find

$$
\left.
\begin{aligned}
\Delta E_{n,l,l+\frac{1}{2},m} &= \frac{lh^2}{16\pi^2\mu^2 c^2} \overline{\left(\frac{1}{r} \frac{dV}{dr} \right)}_{n,l} \\
\Delta E_{n,l,l-\frac{1}{2},m} &= -\frac{(l+1)h^2}{16\pi^2\mu^2 c^2} \overline{\left(\frac{1}{r} \frac{dV}{dr} \right)}_{n,l}
\end{aligned}
\right\}
\quad (61 \cdot 46)
$$

in accordance with our previous result.

If there is a uniform external magnetic field $\vec{\mathcal{K}}$, a term

$$
-\frac{e}{2\mu c} \vec{\mathcal{K}} \cdot (\vec{\mathcal{L}} + 2\vec{s})
$$

should be added to the Hamiltonian of Eq. (61·32). The reader will find it a simple, but instructive, exercise to use this term to work out the theory of the complex Zeeman effect for the alkali atoms.

CHAPTER XIV

THE THEORY OF THE STRUCTURE OF MANY-ELECTRON ATOMS

62. GENERAL FORMULATION OF THE PROBLEM

62a. The Configuration Space.—The most highly developed field of applications of the nonrelativistic quantum mechanics is in the theoretical study of the optical spectra of the elements. The quantum theory had its origin in the attempt to interpret these spectra and gives a highly satisfactory account of their main features. Although the mathematical difficulty of solving the Schrödinger equation for the general many-electron atom is so great that a complete set of exact solutions is wholly unattainable, we can go much farther with available approximations than we can in theoretical studies of the structure of molecules and of the solid state. In this field, moreover, we avoid the basic difficulties which beset the problem of quantum electrodynamics and the problem of nuclear structure. Hence it is appropriate to bring the present volume to a close with an introductory application of our general theory to the central problem of the structure of many-electron atoms and their spectra.

As the wave functions in the Pauli spin theory for a single-electron atom are spread out over a four-dimensional configuration space, the wave functions in an atomic problem involving f electrons require a $4f$-dimensional configuration space. We denote the spin coordinate of the kth electron by m_{sk} or σ_k. There are 2^f possible sets of values of the f spin coordinates and the complete wave function is accordingly the sum of 2^f different $3f$-dimensional "space functions" $u(x_1', y_1', \cdots, z_f')$, each multiplied by a different product of delta functions of the coordinates m_{sk}.

In normalization, the computation of mean values, and the formation of matrix elements, we have to work out scalar products over all configuration space. The process of integration over the $3f$-dimensional positional coordinate space must then be supplemented by a summation over all possible sets of values for the spin coordinates. Using the notation introduced in Sec. 36c we indicate this mixed process of integration by the symbol $\overline{\sum}$. Thus the normalization condition takes the form

$$(\psi,\psi) = \overline{\sum}_{(x_k',\,m_{sk})} |\psi|^2 = \sum_{m_{sk}} \int_\infty |\psi(x_k',m_{sk})|^2 d\tau = 1. \qquad (62\cdot1)$$

523

Functions of the Cartesian space coordinates and the spin coordinates which are of class D in the former for every fixed set of values of the latter are conveniently reckoned as of class D in the combined set of space-spin coordinates, and as physically admissible in the $4f$-dimensional theory. Operators which are Hermitian or unitary with respect to class D and positional coordinate space are also Hermitian, or unitary, as the case may be, with respect to class D and the enlarged $4f$-dimensional configuration space. With the aid of the new coordinate space and the modified definition of class D we are clearly at liberty to carry over bodily the general theory of dynamical variables developed in Sec. 36 into the study of problems involving many particles with spin.

62b. The Hamiltonian Operator.—According to the Kramers classical model of Sec. 58d the secular torque acting on the spin angular momentum of an individual electron is

$$\frac{d\vec{s}}{d\tau} = \frac{e}{\mu c}\vec{s} \times \left\{\vec{\mathfrak{X}} + \frac{1}{2c}[\vec{\mathcal{E}} \times \vec{v}]\right\}. \tag{62·2}$$

The corresponding energy of the spin with respect to the electromagnetic field $\vec{\mathcal{E}}$, $\vec{\mathfrak{X}}$ is

$$U_s = -\frac{e}{\mu c}\vec{s} \cdot \left\{\vec{\mathfrak{X}} + \frac{1}{2c}[\vec{\mathcal{E}} \times \vec{v}]\right\}. \tag{62·3}$$

In the case of a single electron in a Coulomb field, we set $\vec{\mathfrak{X}} = 0$ and $\vec{\mathcal{E}} = -\vec{r}Ze/r^3$. This reduces the spin energy to the form

$$U_s = \frac{Ze^2}{2\mu^2c^2r^3}\vec{s} \cdot \vec{\mathcal{L}} \tag{62·4}$$

given in Eq. (58·6). The expressions for the electric and magnetic forces become much more complicated, however, for many-particle problems. To attempt to take into account the finite speed of propagation of electric and magnetic forces would lead us into the previously mentioned perplexities of quantum electrodynamics. Setting aside these *retardation effects*, as they are called after the retarded potentials used to compute them, and treating the instantaneous field of each electron as if it were made up of a static electric charge e, a steady current element $\vec{ev}$, and a static elementary magnet of moment $e\vec{s}/\mu c$, it is possible to devise a plausible expression for the U_s of a many-electron atom in terms of the space coordinates and momenta of the electrons, the spin operators, and

the potentials of the external electromagnetic field.[1] We shall not undertake here, however, either an examination or application of the complicated Hamiltonian arrived at in this way. The spin energy is in any case relatively small, and the difficulties of an exact treatment of the electrostatic energy are so great as to render somewhat superfluous any attempt to deal very accurately with the spin-orbit and spin-spin energies of the best available Hamiltonian operator.

We are therefore led to seek the simplest approximate expression for the spin energy which will give a qualitative representation of the experimental observations regarding the fine structure of spectrum lines. To this end we neglect the magnetic forces acting on each electron due to the others and identify the field $\vec{\mathcal{E}}$ of Eq. (62·3) with the sum of the external electric field and the average field acting on the electron as a result of the electrostatic forces coming from the nucleus and the other electrons. In other words we set

$$e\vec{\mathcal{E}} = -\frac{\vec{r}}{r}\frac{dV}{dr} + e\vec{\mathcal{E}}'(\vec{r}),$$

where $\vec{\mathcal{E}}'$ is the external electric field and $V(r)$ is identical with the potential of V_0 of Sec. 57b. $\vec{\mathcal{H}}$ is identified with the external magnetic field $\vec{\mathcal{H}}'$ at the point in question. Thus the classical expression (62·3) gives rise to the sum of two terms U_S and U_S' in the Hamiltonian operator, one representing the spin-orbit energy, the other the mutual energy of the spin and the external field. Explicitly these terms are [cf. Eq. (58·5)]

$$U_s = \frac{1}{2\mu^2c^2}\sum_{k=1}^{f}\frac{1}{r_k}\frac{dV(r_k)}{dr_k}\vec{S}_k \cdot \vec{\mathcal{L}}_k, \tag{62.5}$$

$$U_s' = -\frac{e}{\mu c}\sum_{k=1}^{f}\vec{S}_k \cdot \left\{\vec{\mathcal{H}}'(\vec{r}_k) + \frac{1}{2\mu c}[\vec{\mathcal{E}}'(\vec{r}_k) \times \vec{p}_k]\right\}. \tag{62·6}$$

The term U_S is somewhat arbitrary, as $V(r)$ means nothing unless we base the calculation of atomic energy levels on a perturbation computation in which the starting point is the solution of a well-chosen degenerate central-field problem of the type described in Sec. 57b. Since this is about the only possible method of attack on complicated atoms, the use of (62·5) is not unsatisfactory.

---

[1] *Cf.*, *e.g.*, J. FRENKEL, *Wave Mechanics, Advanced General Theory*, section 38, Oxford, 1934. Other treatments of this subject are to be found in the following references: W. Heisenberg, *Zeits. f. Physik* **39**, 499 (1926); W. Pauli, Jr., *Zeits. f. Physik* **43**, 601.(1927); J. A. Gaunt, *Proc. Roy. Soc.* **A122**, 513 (1929), *Trans. Roy. Soc.* **228**, 151 (1929); G. Breit, *Phys. Rev.* **36**, 383 (1930).

In the absence of an external field the complete Hamiltonian for an atom, or atomic ion, neglecting the nuclear motion as in (17·2), becomes

$$H = -\frac{h^2}{8\pi^2\mu}\sum_{k=1}^{f}\nabla_k^2 - \sum_{k=1}^{f}\frac{Ze^2}{r_k} + \sum_{k>j}^{f}\frac{e^2}{r_{kj}} + \frac{1}{2\mu^2 c^2}\sum_{k=1}^{f}\frac{1}{r_k}\frac{dV(r_k)}{dr_k}\vec{S}_k \cdot \vec{\mathfrak{L}}_k.$$

(62·7)

We arbitrarily resolve this Hamiltonian into the sum of three terms as follows:

$$H = H_0 + H_1 + U_s,$$

(62·8)

where

$$H_0 = -\frac{h^2}{8\pi^2\mu}\sum_{k=1}^{f}\nabla_k^2 + \sum_{k=1}^{f}V(r_k),$$

(62·9)

$$H_1 = \sum_{k>j}^{f}\frac{e^2}{r_{kj}} - \sum_{k=1}^{f}\frac{Ze^2}{r_k} - \sum_{k=1}^{f}V(r_k).$$

(62·10)

$V(r)$ is to be so chosen as to make the elements of the matrix $\mathbf{H}_1$ as small as possible. This choice makes H_0 as good an approximation to H as is possible without altering its fundamental form.

62c. The Perturbation Form of the General Atomic Problem.—As explained in Sec. 57b, the usual method of approach to the problem of the many-electron atom is through a perturbation calculation in which the unperturbed problem (i.e., problem A) is of the central-field type with an H_0 of the form specified in Eq. (62·9) permitting solution by the method of the separation of variables. In fact the whole theoretical significance of the empirical analysis of the energy-level systems is tied up with this method of approach. The energy levels of the unperturbed problem are defined by a set of values of the Bohr quantum numbers n and l, one pair for each electron. All energy levels of the perturbed problem originating in a common central-field energy level are said to belong to a common *electronic configuration*. In specifying the individual electron states which define the configuration it is customary to use the spectroscopic term symbols $1s, 2s, 2p, \cdots$, introduced in Sec. 56a. If f electrons have the same pair of values of n and l, the corresponding term symbol is provided with the exponent f. Otherwise the symbol for a configuration is written formally as the product of the term symbols for the individual electrons. For example, the configuration which gives rise to the lowest diffuse-series level for sodium is symbolized by $1s^2 2s^2 2p^6 3d$ to indicate that the core contains two $1s$ electrons, two $2s$ electrons, and six $2p$ electrons, while the valence electron is in a $3d$ state.

The straightforward way to carry through a first-order calculation of the perturbed (true) energy levels would be to form the matrix $\mathbf{H}_1 + \mathbf{U}_s$

with unperturbed wave functions and diagonalize separately the step of this matrix belonging to each individual configuration. This process is difficult, however, owing to the great difference between the symmetry of the unperturbed problem and that of the actual problem. Hence it is desirable, where possible, to make use of the fact that in a large and particularly important class of cases the electrostatic interaction H_1 is much more important than the spin-orbit interaction U_s.

In Sec. 58a we saw that the major energy levels of the lighter atoms, and to a large extent of the heavier atoms as well, can be classified by means of two quantum numbers L and S. The first of these was provisionally interpreted as the quantum number of the resultant orbital angular momentum $\overrightarrow{\mathcal{L}}$ because it obeys the principle of selection $\Delta L = 0$, ± 1. The second determines the multiplicity of the system to which the level belongs and in the lighter atoms is subject to the principle of selection $\Delta S = 0$, which forbids intersystem combination lines. In Sec. 58b we verified that the number of fine-structure components of the various major levels agrees with the number to be expected if S is interpreted as an angular momentum quantum number like L and the corresponding angular momentum vector $\overrightarrow{s}$ is weakly coupled to $\overrightarrow{\mathcal{L}}$. Finally, in Sec. 58d we identified $\overrightarrow{s}$ in the case of the alkalies with the spin angular momentum of the valence electron. It will be evident that in the more general case where there are many electrons outside the inert spherically symmetric core, we can hardly fail to identify $\overrightarrow{s}$ with the resultant spin angular momentum of the valence electrons. In due course we shall prove that the resultant spin angular momentum of the core is zero, so that $\overrightarrow{s}$ can equally well be identified with the resultant spin angular momentum of the entire atom.

The above identifications imply that (a) $\mathcal{L}^2$ and S^2 are approximate integrals of the motion and that (b) they are (in the same approximation) functions of the energy. We know from Secs. 40f and 58b that both of these conditions are satisfied insofar as $\mathcal{L}^2$ is concerned if we neglect the spin energy U_s and spin coordinates. It is easy to see (cf. Sec. 63a) that the introduction of spin coordinates without spin energy does not affect this conclusion. If we neglect H_1 as well as U_s, $\mathcal{L}^2$ remains an integral of the motion, but is no longer a function of the energy since more than one eigenvalue of $\mathcal{L}^2$ is compatible with most eigenvalues of H_0. Therefore the suggested interpretations of the quantum numbers L and S imply that the approximation in which H_1 is taken into consideration, but U_s is neglected, is of primary importance. As a matter of fact this approximation is sufficient in a large class of cases to locate the major energy levels with reasonable precision. The fine structure can

then be calculated by a second perturbation calculation in which U_S is taken into consideration, and the square of the resultant $\vec{\jmath}$ of $\vec{\mathcal{L}}$ and $\vec{S}$ becomes a function of the energy. We owe to Russell and Saunders[1] the first suggestion that the spectra of the alkaline earths can be most easily interpreted on the assumption that in certain energy levels both valence electrons contribute to a quantized resultant orbital angular momentum. Hence it is customary to say that we have to do with *Russell-Saunders coupling* whenever U_S plays a minor role in comparison with H_1. Because the last stage of an energy-level calculation for Russell-Saunders coupling involves what is called "the union of $\vec{\mathcal{L}}$ and $\vec{S}$ to form a quantized resultant $\vec{\jmath}$," Russell-Saunders coupling is also called $\vec{\mathcal{L},\vec{S}}$ coupling. In this book we make no attempt to discuss problems in which the coupling of the electrons is of any other type.

For problems of the class under consideration the perturbation calculation $H_0 \to H_0 + H_1 + U_S$ is conveniently subdivided into two parts, *viz.*, $H_0 \to H_0 + H_1$ and $H_0 + H_1 \to H_0 + H_1 + U_S$. Following the nomenclature of Sec. 57b we designate the H_0 eigenvalue-eigenfunction problem as A, the $H_0 + H_1$ problem as B, and the complete $H_0 + H_1 + U_S$ problem as C. The intermediate Hamiltonian $H_0 + H_1$ will be referred to as H_B.

The division of the complete perturbation calculation $A \to C$ into two parts, $A \to B$ and $B \to C$, is advantageous in providing us with a theoretical interpretation of the empirical L values and their selection rule—also as a labor-saving device. There is an economy of effort because the hypothesis of Russell-Saunders coupling permits us in first approximation to neglect the matrix elements of U_S between different energy levels of problem B, even though these levels have come from the same initial configuration.

63. PROBLEM B: THE SPIN-ORBIT ENERGY NEGLECTED

63a. Integrals of the Motion. —We have learned (Sec. 48d) that in a perturbation calculation it is usually advisable to choose for the initial unperturbed wave functions simultaneous eigenfunctions of the unperturbed Hamiltonian and as many integrals of the perturbed motion as possible. The initial matrix of the perturbed Hamiltonian is then diagonal in the integrals and only one eigenvalue of each integral need be considered at a time. Moreover, any integrals which are functions of the perturbed energy are available for the classification of the perturbed levels.

We have already made a study of the integrals of the spin-free form of problem B in Sec. 40 and have made use of the results at many points

[1] H. N. Russell and F. A. Saunders, *Astrophys. J.* **61**, 38 (1925).

in the last two chapters. The integrals discussed in Sec. 40 belong to two sets associated with the rotation-reflection group and the permutation group, respectively. The addition of the spin coordinates to the positional coordinates brings in the spin operators which are integrals of the perturbed motion originating like $\mathcal{L}_x$, $\mathcal{L}_y$, $\mathcal{L}_z$ in the rotation-reflection group of transformations. On account of the Pauli principle we are interested only in those integrals which commute with the antisymmetrizing operator $\mathcal{G}$ of Sec. 42b and constitute actual observables. This fact rules out the operators for the individual spins and leaves only the components of the resultant spin

$$S_x = \sum_{k=1}^{f} S_{kx}, \qquad S_y = \sum_{k=1}^{f} S_{ky}, \qquad S_z = \sum_{k=1}^{f} S_{kz}, \qquad (63\cdot1)$$

and functions of them. The simplest of these functions which commutes with all three of the primary components is

$$S^2 = S_x{}^2 + S_y{}^2 + S_z{}^2. \qquad (63\cdot2)$$

The vector operator $\vec{S}$ defined by Eqs. (63·1) conforms to the usual commutation rules for angular momenta and is to be identified with the vector $\vec{S}$ of Sec. 58. We shall use the symbols S and M_S for the quantum numbers of S^2 and S_z as in Eq. (58·2). Each eigenvalue of S_z must be the sum of eigenvalues of S_{kz} over all values of k, for the operators S_{kz} commute, and their matrices can be made simultaneously diagonal. Consequently M_S is an integer, or an odd multiple of $\tfrac{1}{2}$, according as the number of electrons attached to the nucleus is even or odd. By the much cited argument of Sec. 40f it follows that S is also an integer, or an odd multiple of $\tfrac{1}{2}$, according as the number of electrons in the atomic system is even or odd. This conclusion gives a theoretical explanation of the empirical observation (p. 495) that energy levels of even multiplicity (even values of $2S + 1$) occur in atoms with an odd number of extranuclear electrons, while those of odd multiplicity occur in atoms with an even number of extranuclear electrons.

We saw in Sec. 40f, p. 314, that $\mathcal{L}^2$ and K commute with every member of the group of the Schrödinger equation for the spin-free form of problem B and are therefore implicit functions of the Hamiltonian in the discrete spectrum. Since the spin operators all commute with $\mathcal{L}^2$ and K this conclusion is not affected by the introduction of the spin coordinates. (In fact we have already assumed in Sec. 62c that $\mathcal{L}^2$ is a function of the energy in problem B.) It follows that every energy level of problem B is characterized by a definite value of the quantum number L and has $2L + 1$ mutually orthogonal eigenfunctions which are also eigenfunctions

of $\mathcal{L}_z$ with different eigenvalues and which are obtainable one from the other by the application of the operators $\mathcal{L}_x \pm i\mathcal{L}_y$.

S^2 also commutes with every integral of problem B, nonsymmetric functions of the individual spin operators S_{kx}, S_{ky}, S_{kz} excepted. Since the nonsymmetric spin operators do not commute with the antisymmetrizing operator $\mathcal{G}$ of Sec. 42b, we can say that S^2 commutes with every operator which commutes with H_B and $\mathcal{G}$. Consequently S^2 is a function of H_B and $\mathcal{G}$. But by the Pauli principle *every* physically admissible wave function is an eigenfunction of $\mathcal{G}$ with the eigenvalue $+1$. This means that for physical purposes S^2 may be regarded as a function of H_B, for, if an atom is in a state whose wave function is a physically admissible eigenfunction of H_B, it will be correlated with a definite value of S.

This conclusion is at first sight very surprising, especially when we verify that a single configuration and a single value of $\mathcal{L}^2$ can yield two different and well-separated energy levels of problem B with different values of S^2. When it was first discovered experimentally that atomic energies had an apparently strong dependency on the associated values of S, it seemed necessary to assume a strong spin-spin interaction energy to account for the phenomenon. Our analysis indicates, however, that this strong dependency of energy on the resultant spin quantum number can exist without any spin-energy term in the Hamiltonian whatsoever. Actually the energy differences of states with different values of S, but otherwise apparently the same, are of electrostatic origin and simply reflect different types of symmetry in the wave functions.

Because of the $2S + 1$ values of M_S compatible with the quantum number S, every energy level associated with this quantum number has $2S + 1$ mutually orthogonal wave functions derivable one from the other by application of the operators $S_x \pm iS_y$. Since these operators commute with $\mathcal{L}_z$ the total number of mutually orthogonal wave functions of an energy level with the orbital- and spin-angular-momentum quantum numbers L and S is at least $(2L + 1)(2S + 1)$, or the product of the number of M_L values compatible with L into the number of M_S values compatible with S.

The energy levels of problem B, hitherto called "major energy levels," are sometimes called "terms" to distinguish them from the energy levels of the central-field problem A (electronic configurations) and from the final fine-structure components of problem C.[1] They are classified by the corresponding eigenvalues of $\mathcal{L}^2$, S^2, and K. The letter symbols indicating different values of the quantum number L are indicated in Table I, Sec. 58a. The eigenvalues of S^2 are indicated by an appropriate numerical superscript—the multiplicity $2S + 1$ of the corresponding system—attached to the letter symbol on the left. Thus ^3S and ^2P

•   [1] *Cf.* Condon and Shortley *T.A.P.*, p. 189.

denote, respectively, a triplet term ($S = 1$) with L equal to zero and a doublet term ($S = \frac{1}{2}$) with L equal to unity. In order to distinguish between the even and odd states correlated with the eigenvalues $+1$ and -1 of the operator K, a superscript 0 may be attached on the right to the term symbols of odd states. Thus D^0 denotes an odd state for which $L = 2$ and D an even state with the same value of L.

As a matter of fact, K is a function of the Hamiltonian H_0, so that all energy levels coming from any given configuration are either even, or odd, as the case may be. To see which configurations give even levels, and which odd, we apply the operator K first to the individual solutions of the central-field problem, $\psi_{n,l,m}$, obtained by separating the variables in spherical coordinates r, θ, φ. When applied to functions of r, θ, φ, the operator K replaces θ by $\pi - \theta$ and φ by $\pi + \varphi$. The Legendre polynomials $P_l(\cos \theta)$ (*cf.* Sec. 27b) are even or odd functions of $\cos \theta$ according as l is even or odd. Hence the associated Legendre functions $P_{l,|m|}(\cos \theta)$ are even or odd functions of $\cos \theta$ according as $l - m$ is even or odd. Thus K reverses the sign of $P_{l,|m|}(\cos \theta)$, or leaves it unchanged, according as $l - m$ is even or odd. But $K e^{im\varphi} = e^{im\varphi}$ if m is even and

$$K e^{im\varphi} = -e^{im\varphi}$$

if m is odd. Thus $\psi_{n,l,m}$ is even or odd with respect to K according as l is even or odd, independent of m.

Let us next consider a product-form eigenfunction of the many-electron atomic problem in the limiting case A form. Evidently the effect of K is to introduce a factor $(-1)^{l_k}$ for each electron. Hence *a configuration is even or odd according as $\sum_{k=1}^{f} l_k$ is even or odd.*

In considering the problem of degeneracy for a Hamiltonian operator H in the light of the Pauli exclusion principle it is necessary to distinguish between two kinds of degeneracy. The *mathematical* degeneracy of an energy level E_k is the total number of linearly independent eigenfunctions of E_k, regardless of sex, race, or previous condition of servitude. The discussion in Secs. 40e and f refers to this kind of degeneracy. On the other hand we define the *physical* degeneracy of E_k as the number of linearly independent *antisymmetric* eigenfunctions associated with E_k, provided, of course, that the system under consideration contains at least two electrons. Physical and mathematical degeneracy are to be identified for systems that contain no pairs of identical particles. The physical degeneracy gives the statistical weight g_k introduced in Sec. 54a for the calculation of population densities, since only antisymmetric wave functions can support a population.

In connection with our study of the Pauli principle in Sec. 42b we defined a complete set of normally commuting independent observables

as a set with the property that any pair of antisymmetric simultaneous eigenfunctions of the set with the same eigenvalues must be linearly dependent. In other words, the sets of mutually compatible eigenvalues of such a set of observables are all to be *physically nondegenerate*. The complete set of commuting observables bears the same relation to the problem of physical degeneracy as the complete set of commuting dynamical variables bears to the problem of mathematical degeneracy. If the Hamiltonian H is united with observables α_1, α_2, $\cdots$ to form a complete set of commuting observables, the physical degeneracy of any energy level E_k is equal to the number of different sets of α eigenvalues compatible with E_k.

In view of the spin degeneracy it is obvious that the set of independent commuting observables formed by H_B and $\mathcal{L}_z$ is *not* complete. We shall prove that it becomes complete when we add S_z to it. It is to be observed that the permutation group contributes nothing to the physical degeneracy (*cf.* Sec. 42b, p. 338). Thus the factor N in the expression $N(2L + 1)$ given on p. 317 for the number of linearly independent simultaneous wave functions belonging to E and L, becomes unity if we count only functions which satisfy the Pauli principle. On the other hand, the inclusion of the spin coordinates in the wave functions enlarges the rotation-reflection group of operators. It is possible to rotate the reference axes for the spin coordinates and for the space coordinates *separately*. Hence the physical degeneracy of an energy level E_k with the orbital angular-momentum quantum number L and the spin quantum number S is equal to the product of the number of wave functions derivable from any one by rotations of the axes for the space coordinates and the number derivable by rotations of the axes for the spin coordinates.

In Sec. 40c we saw that every rotation of the space-coordinate axes can be effected by an operator of the form

$$F = e^{\frac{2\pi i}{h}\omega(\lambda\mathcal{L}_x + \mu\mathcal{L}_y + \nu\mathcal{L}_z)}, \tag{63.3}$$

where λ, μ, ν are direction cosines. Such an operator can commute with $\mathcal{L}_z$ only if λ and μ are zero, in which case F becomes a function of $\mathcal{L}_z$.

Inspection shows that the operator $T(\omega)$ which transforms a spin function from one set of reference axes to another obtained from it by rotation through an angle ω about the z axis is of the form

$$T(\omega) = e^{\frac{2\pi i}{h}\omega S_z} \tag{63.4}$$

(*cf.* Sec. 61b). The corresponding operator for a general rotation about an axis whose direction cosines are λ, μ, ν is accordingly

$$T(\vec{\omega}) = e^{\frac{2\pi i}{h}\omega(\lambda S_x + \mu S_y + \nu S_z)} \tag{63.5}$$

Evidently the only operators of this type which can commute with S_z are functions of S_z.

To complete our discussion of the rotation-reflection group for the spin variables it is natural to attempt to set up a spin operator analogous to the reflection operator K. An examination of the problem shows, however, that it is impossible to construct a spin operator whose properties offer a satisfactory parallel to those of K. Even if it could be done we should expect the new operator to commute with all the rotations and hence to be a function of S^2 of little importance.

It is now evident that there is no pair of observables in the group of the Schrödinger equation which commute with H_B, $\mathcal{L}_z$, S_z, but not with each other. In accordance with the fundamental postulate of Sec. 40e, p. 313, we conclude that these three operators form (within the discrete spectrum of H_B) a complete set of commuting independent observables and that the statistical weight of an energy level with the orbital- and spin-angular-momentum quantum numbers L and S is exactly

$$(2L + 1)(2S + 1)$$

as indicated on p. 530.

63b. Antisymmetric Functions and the Empirical Pauli Exclusion Rule.—The simplest class of eigenfunctions of the Hamiltonian H_0 of of problem A are the products of individual electron functions similar to those introduced in Eq. (57 6), except that they contain the spin coordinates. Each of the individual electron functions u has for its arguments the coordinates x', y', z', σ of some particular electron, say the kth, and is a solution of the corresponding equation

$$\nabla_k^2 u + \kappa[\epsilon - V(r_k)]u = 0 \tag{63·6}$$

[cf. Eq. (57·5)]. Solutions of this equation obtained by the usual separation of variables are characterized by a set of four individual electron quantum numbers n, l, m_l, m_s. Let the a_j denote the jth of these individual sets and let $(k|a_j)$ denote the corresponding solution of (63·6). It is the product of a function of the form of ψ_{n,l,m_l} defined in Eq. (57·7) and a spin factor δ_{σ_k,m_s}. Denoting a set of f sets of four individual electron quantum numbers by α, we introduce

$$\phi_\alpha = (1|a_1)(2|a_2) \cdots (f|a_f) = \prod_{k=1}^{f} (k|a_k) \tag{63·7}$$

as the typical product-form eigenfunction of H_0 with the eigenvalue

$$E_0(\alpha) = \sum_{j=1}^{f} \epsilon(a_j).$$

The energy $E_0(\alpha)$ is degenerate on account of the multiplicity of m_l and m_s values which can be assigned to any particular pair of values of n

and l. There is also degeneracy due to the equivalence of the different functions obtained by applying the various permutation operators to ϕ_α. The latter degeneracy is eliminated, however, by the Pauli principle which admits at most one function obtained from ϕ_α by permutation and linear combination, *viz.*, the function $\mathcal{G}\phi_\alpha$ obtained from ϕ_α by application of the operator $\mathcal{G}$.

$\mathcal{G}\phi_\alpha$ *will vanish if, and only if, the full set of 4f quantum numbers α includes two or more identical individual sets of four.* To prove this statement we note that if two sets of quantum numbers j and j' are equal, each term of the sum

$$\mathcal{G}\phi_\alpha = \frac{1}{f!}\sum_{\tau=1}^{\tau=f!}(-1)^{p_\tau}P_\tau\phi_\alpha \qquad . \qquad (63\cdot8)$$

can be paired with another term which differs from it only in the interchange of the coordinates of the two electrons which in these terms have the same set of quantum numbers. The permutations which give these terms differ by a simple interchange. Hence one is odd and the other even. It follows that the terms are equal in magnitude and opposite in sign. Hence all terms cancel in pairs when two or more individual sets of quantum numbers are equal. If no two sets are equal, the permutations of ϕ_α are linearly independent, and the linear combination $\mathcal{G}\phi_\alpha$ cannot vanish.

We conclude that in the central-field approximation *no states are allowed by the wave-mechanical form of the Pauli principle developed in Sec. 42b in which two electrons have the same set of four quantum numbers n, l, m_l, m_s.* Since m_s has just *two* values, this means that not more than *two* electrons can have the same set of three space-coordinate quantum numbers n, l, m_l. Thus the wave-mechanical form of the Pauli exclusion principle is actually equivalent to the more empirical exclusion rule given in Sec. 57b.

Because $\mathcal{G}$ commutes with H_0 and H_1 we know from Sec. 48d that in solving the perturbation problem $A \rightarrow C$ we can deal separately with each eigenvalue of $\mathcal{G}$. As λ approaches zero each antisymmetric eigenfunction of $H_0 + \lambda H_1$ must approach an antisymmetric eigenfunction of H_0. The interpolation problem will establish a connection between each antisymmetric eigenfunction of problem B and a corresponding antisymmetric eigenfunction of problem A. Therefore the exclusion of problem A wave functions which are not antisymmetric implies the exclusion of the problem B wave functions associated with them.

The reader will recall that in the discussion of the equivalence of particles of the same species in Sec. 42b we were able to prove theoretically that physically admissible wave functions must be *either symmetric or*

antisymmetric with respect to all interchange permutations P_{ij} involving a single kind of particle. In order to complete the formulation of the Pauli principle it was necessary, however, to make the arbitrary assumption that the symmetric possibility is to be ruled out for electrons and protons, physically admissible ψ functions being definitely antisymmetric with respect to interchanges of these two species. We are now in a position to justify this assumption insofar as it applies to electrons.

Symmetric eigenfunctions of H_0 exist with any number of identical sets of individual quantum numbers n, l, m_l, m_s. Consequently the use of symmetric wave functions would *not* involve an exclusion rule of the type demanded by the empirical facts cited in Sec. 57b. Therefore the above mentioned assumption is required by the experiment. The corresponding proof that physically admissible wave functions must also be antisymmetric with respect to proton interchanges is obtainable in essentially the same way, *viz.*, by a comparison of the observed energy levels of a multiproton problem—that of the H_2 molecule—with those to be expected from the symmetric and antisymmetric alternatives allowed by the general theory.[1]

Although the application of the operator $\mathcal{G}$ to the product-form eigenfunction of problem A given in (63·7) generates an antisymmetric eigenfunction of the same problem, it does not take care of the normalization of the new function. Let us assume that ϕ_α is normalized to unity by the normalization of the factors $(k|a_k)$. The scalar product of any pair of terms in $\mathcal{G}\phi_\alpha$ resolves itself into a product of individual scalar products such as

$$((k|a_j),(k|a_{j'})) = \sum_{\sigma_k} \int \int \int (k|a_j)(k|a_{j'})^* dx_k' dy_k' dz_k'.$$

If $\mathcal{G}\phi_\alpha$ does not vanish identically, at least two of the factors in any term will be orthogonal to the corresponding factors of any other specific term. Hence every term is orthogonal to every other, and

$$(\mathcal{G}\phi_\alpha, \mathcal{G}\phi_\alpha) = \left(\frac{1}{f!}\right)^2 \sum_{\tau=1}^{f!} (P_\tau\phi_\alpha, P_\tau\phi_\alpha) = \frac{1}{f!}. \qquad (63·9)$$

To obtain normalized functions which satisfy the Pauli principle we multiply $\mathcal{G}\phi_\alpha$ by $\sqrt{f!}$ and call the result $\hat{\mathcal{G}}\phi_\alpha$ or Φ_α. The operator $\hat{\mathcal{G}} = \sqrt{f!}\mathcal{G}$ has been called the *antisymmetrizer* by Condon and Shortley.[2]

It was first pointed out by Slater[3] that the function Φ_α can be written as a determinant built up from the functions $(k|a_j)$. Thus

[1] *Cf.* D. M. DENNISON, *Proc. Roy. Soc.* **A115**, 483 (1927).
[2] E. U. CONDON and G. H. SHORTLEY, *T.A.S.*, p. 165.
[3] J. C. SLATER, *Phys. Rev.* **34**, 1293 (1929).

$$\Phi_\alpha = \hat{\mathsf{G}}\phi_\alpha = \frac{1}{\sqrt{f!}} \begin{vmatrix} (1|a_1) & (2|a_1) & \dots & (f|a_1) \\ (1|a_2) & (2|a_2) & \dots & (f|a_2) \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ (1|a_f) & (2|a_f) & \dots & (f|a_f) \end{vmatrix} \qquad (63\cdot10)$$

The equivalence of (63·10) and (63·8) follows directly from the definition of a determinant. Wave functions of this form are commonly referred to as Slater wave functions. When we express $\hat{\mathsf{G}}\phi_\alpha$ as a determinant, its antisymmetry with respect to electron interchanges follows from the elementary rule that a determinant changes sign whenever any two columns are interchanged.

63c. Closed Shells.—A group of electrons to which the same values of n and l are assigned in the case A approximation is said to constitute a "shell."[1] The electrons in any given shell are said to be *equivalent*. By the Pauli exclusion rule the maximum number of electrons which can coexist in such a shell is equal to the number of different pairs of values of m_l and m_s which are compatible with l. This number is $2(2l + 1)$. A shell containing the maximum allowed number of electrons is said to be *closed*.

The complete set of individual sets of quantum numbers required for a closed shell is uniquely defined. Hence the corresponding wave function Φ is uniquely defined except for the usual arbitrary phase factor. The quantities $M_L = \sum_{k=1}^{f} m_{lk}$ and $M_S = \sum_{k=1}^{f} m_{sk}$ are both zero for every term of Φ. It follows that Φ is a simultaneous eigenfunction of $\mathcal{L}_z$, S_z, J_z with the eigenvalue zero for each. It can also be proved to be an eigenfunction of $\mathcal{L}^2$, S^2, J^2, with the common eigenvalue zero. To establish this proposition we recollect that according to Sec. 63a $\mathcal{L}^2$, $\mathcal{L}_z$, S^2, S_z, and $\hat{\mathsf{G}}$ are integrals of the motion for problem B—and *a fortiori* for problem A. Since they commute, it is possible to make their matrices and that of H_0 simultaneously diagonal. Hence it must be possible to construct simultaneous eigenfunctions of all of these operators from Φ and other physically admissible wave functions for the same shell. But this shell does not have two linearly independent physically admissible wave functions. Hence Φ must be from the beginning an eigenfunction of $\mathcal{L}^2$ and S^2. The only eigenvalue of either of these observables compatible with the lack of degeneracy is zero. But the eigenvalue of $\mathcal{L}^2$

---

[1] The notation goes back to the static atomic models of G. N. Lewis and Irving Langmuir in which the electrons were thought of as arranged in successive symmetric groups, or "shells," surrounding the nucleus.

for any given eigenfunction is the sum of the mean values of $\mathcal{L}_x{}^2$, $\mathcal{L}_y{}^2$, $\mathcal{L}_z{}^2$ for the state in question. These mean values are all positive or zero. It follows that Φ is a simultaneous eigenfunction of $\mathcal{L}_x$, $\mathcal{L}_y$, $\mathcal{L}_z$ with the eigenvalue zero for each. In the same way we see that each of the components of $\vec{S}$ yield zero when applied to Φ. Finally, it follows that each component of $\vec{\mathcal{J}}$ reduces Φ to zero identically and that Φ is an eigenfunction of $\mathcal{J}^2$ with the eigenvalue zero.

Since the resultant orbital, spin, and total angular momenta of a closed shell are all zero, we are led to expect such a shell to be spherically symmetric. This expectation is verified by the observation that the operators F and T of Eqs. (63·3) and (63·4) which rotate the reference axes for the space and spin coordinates, respectively, through arbitrary angles, both reduce to the identical operator when applied to the wave function of a closed shell.

In passing we note that a rotation of common reference axes for space and spin coordinates, such as we ordinarily employ, is effected by the operator $e^{\frac{2\pi i}{h}\omega(\lambda\mathcal{J}_x + \mu\mathcal{J}_y + \nu\mathcal{J}_z)}$ Hence it is clear that any eigenfunction of $\mathcal{J}^2$ with the eigenvalue zero is spherically symmetric. Similarly a spin-free eigenfunction of $\mathcal{L}^2$ is spherically symmetric.

63d. Terms Originating in a Given Configuration.—The perturbation problem $A \rightarrow B$ traces the coupling of the individual orbital angular momenta $\vec{\mathcal{L}}_k$ into a resultant $\vec{\mathcal{L}}$, and that of the individual spin angular momenta $\vec{S}_k$ into a resultant $\vec{S}$. As a consequence of this coupling the configurations of problem A are resolved into groups of terms, each partially, or wholly, characterized by the appropriate values of the resultant orbital and spin quantum numbers L and S. We have now to consider just which term types, and how many of each, originate in any given configuration.

Let us first deal with the coupling of two sets of electrons with the resultant angular-momentum quantum numbers L_1, S_1 and L_2, S_2, respectively. The problem is analogous to that of Sec. 58b but differs from it on account of the inclusion of spin coordinates and the spin angular momenta. We suppose that the electrons of each set are initially coupled together, so that each energy level of each set has definite corresponding values of L and S. The electrons of each set when uncoupled will pass into problem A configurations characterized by a corresponding number of pairs of n,l values. We choose for our consideration the coupling of states of the two sets of such a character that no pair of n,l values which enters into the description of the configuration of one set is repeated in the description of the configuration of the other.

In other words the two sets are to include no two equivalent electrons. By this restriction we avoid the exclusion of any set of values of M_{L1}, M_{S1}, M_{L2}, M_{S2} on account of the Pauli principle.

The argument of Sec. 58b can now be used to demonstrate that for any given values of M_{S1} and M_{S2} the construction of eigenfunctions of $\mathcal{L}^2$ from the various possible wave functions of the uncoupled system will yield values of L ranging from $L_1 + L_2$ down to $|L_1 - L_2|$ with no duplications. For each of these values of L there will be a corresponding L complex of simultaneous eigenfunctions of $\mathcal{L}^2$ and $\mathcal{L}_z$. In the same way linear combinations of wave functions with fixed values of L and M_L $(= M_{L1} + M_{L2})$ but various values of M_{S1} and M_{S2} yield simultaneous eigenfunctions of S^2 and S_z in which S ranges from $S_1 + S_2$ down to $|S_1 - S_2|$, while M_S ranges from $-S$ to $+S$. In other words the L values obtained in this way are the same as those derived in Sec. 58b; the S values are derived by the same rule; and all combinations of L and S values occur. Suppose, for example, that $L_1 = L_2 = 1$, while $S_1 = \frac{1}{2}$ and $S_2 = 1$. The coupling of the two sets of electrons yields the term types ^2S, ^2P, ^2D, ^4S, ^4P, ^4D.

An immediate corollary on the above analysis is that the coupling of a closed shell with a second group of electrons which belong to another shell or shells does not affect the angular momentum of the latter group or the degeneracy of the energy levels. The ordinary optical atomic-energy levels originate in configurations in which all but a few valence .electrons are in closed shells. Hence it is unnecessary to consider any but the valence electrons in determining the number of terms of various types which originate in any given configuration.

The electrons outside the closed inner shells may be concentrated in a single outer shell, or distributed over several such shells. If we first determine the L and S values which originate in the coupling of the electrons in each of the outer shells, the L and S values for the entire atom can be worked out by a succession of imaginary coupling processes in each of which two groups of nonequivalent electrons are united, i.e., by a repetition of the process we have just described. When only two groups of electrons are coupled, a given pair of initial states yields not more than one term of a given type, but this is not true when more than two shells contribute to the angular momentum. The way in which this happens and the method used for labeling the terms in such a case are best explained by an example.

Consider three nonequivalent p electrons, say $3p$, $4p$, $5p$. The coupling of the $3p$ and $4p$ electrons gives rise to terms in which L ranges from 0 to 2 and S takes on the values 0 and 1. The corresponding term types are ^1S, ^1P, ^1D, ^3S, ^3P, ^3D. In coupling the $5p$ electron to the other pair we have to consider the union of each of the above intermediate terms with the new electron. We obtain the following results

$$3p4p \ ^1S + 5p \to \ ^2P,$$
$$3p4p \ ^1P + 5p \to \ ^2S, \ ^2P, \ ^2D,$$
$$3p4p \ \ D + 5p \to \ ^2P, \ ^2D, \ ^2F,$$
$$3p4p \ ^3S + 5p \to \ ^2P, \ ^4P,$$
$$3p4p \ ^3P + 5p \to \ ^2S, \ ^2P, \ ^2D, \ ^4S, \ ^4P, \ ^4D,$$
$$3p4p \ ^3D + 5p \to \ ^2P, \ ^{.2}D, \ ^2F, \ ^4P, \ ^4D, \ ^4F.$$

In this case we have as many as six different terms of the same type as regards L and S values originating in a single configuration of three electrons. In order to distinguish between the different terms of the same type it is necessary to specify the parent term due to the preliminary coupling of the $3p$ and $4p$ electrons from which the given term issues. Thus the four different ^2D terms are indicated by the symbols $3p4p \ (^1P) \ 5p \ ^2D$; $3p4p \ (^1D) \ 5p \ ^2D$; $3p4p \ (^3P) \ 5p \ ^2D$; $3p4p \ (^3D) \ 5p \ ^2D$.

The problem of the coupling of the equivalent electrons remains to be considered. In order to find the terms generated it is necessary to write out in detail the possible sets of m_l and m_s values consistent with the given configuration and, after eliminating those which violate the Pauli principle, to determine by the counting process of Sec. 58b the character of the possible L and S complexes to be obtained by linear combination of the wave unctions associated with the remaining sets.

The procedure is again conveniently indicated by an example. We choose the simple case of two equivalent p electrons—a p^2 configuration. Each wave function is then specified by a corresponding set of values of $m_{l1}, \ m_{s1}, \ m_{l2}, \ m_{s2}$. We adopt the notation of Condon and Shortley,[1] denoting the eigenvalues of the spin coordinates by plus and minus signs attached as superscripts to the corresponding numerical value of $\cdot m_l$. Thus $(1^+ \ -1^-)$ denotes a state, or wave function, derived by application of the antisymmetrizer $\widehat{\mathsf{G}}$ to a product function $u_1(1)u_2(2)$ in which $m_{l1} = 1, \ m_{s1} = \frac{1}{2}, \ m_{l2} = -1, \ m_{s2} = -\frac{1}{2}$. Eliminating all such states in which the two pairs of quantum numbers are the same, and arranging the resulting state symbols in rows and columns according to the resulting values of $M_L = m_{l1} + m_{l2}$ and $M_S = m_{s1} + m_{s2}$, we obtain Table III shown on p. 540. The maximum value of M_L is 2 and the only associated value of M_S is zero. This indicates a ^1D state. The maximum value of M_S is unity and the accompanying values of M_L are $+1, 0, -1$. These indicate a ^3P state. Striking out one symbol from each square of the center column for the ^1D state and one symbol from each square of each of the three central rows for the ^3P state, there remains only a single symbol in the center of the diagram for a ^1S state. Hence the p^2 configuration yields the L and S complexes for ^1S, ^1D, and ^3P terms with no wave functions left over.

[1] CONDON and SHORTLEY, *T.A.S.*, p. 169.

Tables of this type can be constructed for any shell populated with any allowed number of electrons, but increase in complexity with the l value for the shell and up to a certain point with the number of electrons in the shell. Only one wave function and one term type come from a closed shell, however, and it is easy to see that the complexity of the table of wave functions must decrease as the shell approaches completion. Consider the case of a group of four equivalent p electrons, which differs from a closed p shell by the absence of two members of its

TABLE III

		M_S		
		1	0	-1
	2		$(1^+ 1^-)$	
	1	$(1^+ \quad 0^+)$	$(1^+ 0^-)(1^- 0^+)$	$(1^- \quad 0^-)$
M_L	0	$(1^+ -1^+)$	$(1^+ -1^-)(0^+ 0^-)(1^- -1^+)$	$(1^- -1^-)$
	-1	$(0^+ -1^+)$	$(0^+ -1^-)(0^- -1^+)$	$(0^- -1^-)$
	-2		$(-1^+ -1^-)$	

full complement. There are six possible pairs of values of m_l and m_s for p electrons. Each way of picking out n of the six pairs defines a product wave function for n equivalent p electrons. Every scheme for picking out four pairs is also a scheme for picking out two pairs for the two missing electrons. Since the values of $\sum_k m_{lk}$ and $\sum_k m_{sk}$ for the missing electrons are equal and opposite to the values of the same quantities for the electrons actually present, the table of M_L and M_S values for the possible wave functions of the missing electrons is the same as that for the four electrons present. Consequently the terms originating in a configuration of 4 or, more generally, n equivalent p electrons are the same as those originating in a configuration of $6 - 4$, or $6 - n$ equivalent p electrons, as the case may be. This fact greatly simplifies the labor of working out the term types for partially filled shells of equivalent electrons.[1]

64. SELECTION RULES FOR ELECTRIC DIPOLE RADIATION

64a. The Laporte Rule.—The simplest to derive of all the selection rules is the Laporte rule, which states that *dipole-radiation transitions*

[1] The reader is referred to Condon and Shortley, *T.A.S.*, p. 208, for a table of the term types originating in shells of equivalent s, p, d, f electrons in all stages of occupancy.

occur between even states and odd states, but never between even states and even, or between odd states and odd.

Consider the matrix element

$$\vec{D}(A,B) = (\vec{D}\psi_B, \psi_A),$$

in which ψ_A and ψ_B are two eigenfunctions of the Hamiltonian of a many-particle problem. If the integrand $\psi_A{}^*\vec{D}\psi_B$ is antisymmetric with respect to the triple-reflection operator K, it means that for every point of configuration space there is another point, derived from the first by reversing the signs of all the space coordinates, at which the function has the same value except for a reversal of sign. Hence the contributions of the different volume elements cancel in pairs and the sum-integral reduces to zero. But the electric moment $\vec{D}$ is reversed in sign by K. Hence the product $\psi_A{}^*\vec{D}\psi_B$ is antisymmetric if ψ_A and ψ_B have the same symmetry with respect to K. It follows that the matrix elements of $\vec{D}$ between states of the same symmetry with respect to K are all zero, and that no transitions between such states occur as a result of dipole radiation. This rule holds equally well, whether or not we take into account the perturbative terms H_1 and U_s in the Hamiltonian. The rule regarding transitions due to magnetic dipole and electric quadrupole radiation is just the opposite, however, since these quantities are invariant with respect to K.

64b. Selection Rules for the Central-field Problem.—Selection rules for the central-field problem A are of interest because they have approximate validity for the unsimplified problem C in the large class of cases in which the central-field approximation is a fairly good one. In Sec. 55 we derived the rules $\Delta m = 0, \pm 1; \Delta l = \pm 1$ for dipole radiation in the two-particle central-field problem. We shall now prove that these rules are directly applicable to many-electron atoms in the case A approximation.

Let $\Phi_\alpha = \hat{G}\phi_\alpha$ and $\Phi_\beta = \hat{G}\phi_\beta$ denote two normalized central-field wave functions of the kind defined by Eqs. (63·8) and (63·10). The matrix element

$$D_x(\alpha,\beta) = (D_x\Phi_\beta, \Phi_\alpha) \tag{64.1}$$

expands into the double sum

$$D_x(\alpha,\beta) = \frac{1}{f!}\sum_{\lambda=1}^{f!}\sum_{\mu=1}^{f!}(-1)^{p_\lambda+p_\mu}(D_xP_\lambda\phi_\beta, P_\mu\phi_\alpha).$$

Each term of this sum is invariant of a permutation of the electron coordi-

nates in the integrand.   Moreover, D_x is independent of every permutation.   We can accordingly replace $(D_x P_\lambda \phi_\beta,\ P_\mu \phi_\alpha)$ by $(D_x P_\mu{}^{-1} P_\lambda \phi_\beta,\ \phi_\alpha)$, where $P_\mu{}^{-1}$ is the inverse of P_μ.   The products $P_\mu{}^{-1} P_\lambda$ are, of course, members of the permutation group.   The set of products obtained by holding either λ or μ fast and allowing the other to range through its allowed set of $f!$ values contains all members of the group with no duplications.   Hence every permutation P_τ is contained exactly $f!$ times in the complete two-dimensional array of products $P_\mu{}^{-1} P_\lambda$ involved in the above expression for $D_x(\alpha,\beta)$.   Furthermore, it is easy to see that if $P_\mu{}^{-1} P_\lambda = P_\tau$, it follows that $(-1)^{p_\lambda + p_\mu} = (-1)^{p_\tau}$.   Hence

$$
\begin{aligned}
D_x(\alpha,\beta) &= \sum_{\tau=1}^{f!} (-1)^{p_\tau} (D_x P_\tau \phi_\beta,\ \phi_\alpha) \\
&= e \sum_{k=1}^{f} \sum_{\tau=1}^{f!} (-1)^{p_\tau} (x_k' P_\tau \phi_\beta,\ \phi_\alpha).
\end{aligned}
\tag{64·2}
$$

Due to the orthogonality of every pair of functions $(k|a_j)$, $(k|a_{j'})$ for which the two sets of quantum numbers a_j, $a_{j'}$ are not equal, the scalar product $(x_k' P_\tau \phi_\beta,\ \phi_\alpha)$ will vanish if any factor of ϕ_α, except the one containing the coordinates of the kth electron, differs from the corresponding factor of $P_\tau \phi_\beta$ (we take corresponding factors to be those involving the coordinates of the same electron).   This means that, if $D_x(\alpha,\beta)$ is not to vanish, each of the individual sets of four quantum numbers n, l, m_l, m_s used in defining ϕ_α, with one exception, must be paired with an identical set used in defining ϕ_β.   In other words it must be possible to transform ϕ_β by means of a properly chosen permutation into a form which is identical with that of ϕ_α except for a single factor.   A transition between states Φ_α and Φ_β, which differ in this way only, is called a *single-electron jump*.   Since the result holds also for the elements of $\mathbf{D}_y$ and $\mathbf{D}_z$, we conclude that *in the central-field approximation A the only allowed transitions are those of the single-electron type.*

We now assume that the above condition is satisfied.   If $(x_k' P_\tau \phi_\beta,\ \phi_\alpha)$ is not to vanish, k must be the ordinal number of that factor in ϕ_α which does not match the corresponding factor of $P_\tau \phi_\beta$.   Also the permutation P_τ is uniquely determined   Hence the sum on the right of (64·2) reduces to a single term.   By proper choice of the order of the individual sets of quantum numbers b_1, b_2, $\cdots$, b_f which define ϕ_β, we can insure that P_τ shall reduce to the identical permutation.   Let us assume that the required choice has been made.   Equation (64·2) becomes

$$
D_x(\alpha,\beta) = e(x\ '(k|b_k),\ (k|a_k)).
$$

Thus the matrix element of D_x (or D_y, or D_z) for an allowed problem-A transition is identical with the corresponding element of the matrix for

a single electron. It follows that in the central-field approximation *the selection rules and intensities for the many-electron atom are the same as for the single-electron problem with the same potential-energy function,* viz.,

$$\Delta l_k = \pm 1, \qquad \Delta m_{lk} = 0, \pm 1.$$

This conclusion substantiates the assumption of Sec. 57*b*, p. 488.

64c. Selection Rules for Problems B and C.—Selection rules for problem *B* are of interest for the same reason as those for the central-field problem, *i.e.*, because of their approximate validity for the unmodified problem *C*. They are of course more widely applicable to actual atoms than the central-field rules because the Hamiltonian of *B* is better than that of *A*. In order to work out the rules for problem *B* we have to determine which matrix elements of the electric moment vanish when that matrix is worked out from a scheme of wave functions which makes L^2 and S^2 diagonal. It is convenient to assume that L_z, S_z are also diagonal, giving what is known as an LM_LSM_S scheme of wave functions.

The first selection rule to be considered is that which forbids the occurrence of intersystem combination lines, *i.e.*, transitions involving changes in the quantum number S. For this purpose it suffices to note that the product of any eigenfunction of S^2 with a function of the space coordinates only, such as D_x, D_y, D_z, is also an eigenfunction of S^2 with the same eigenvalue as the original function. It follows that $(\vec{D}\psi_B, \psi_A)$ vanishes if ψ_A and ψ_B are eigenfunctions of S^2 with different values of S. This proves the selection rule, which holds very well experimentally for the lighter atoms.

The selection rules for M_L and L are

$$\Delta M_L = 0, \pm 1, \tag{64·3}$$
$$\Delta L = 0, \pm 1. \tag{64·4}$$

The latter rule was used in Sec. 58*a*, p. 493. Closely related to the above, and derivable by the same procedure, are the rules

$$\Delta M = 0, \pm 1, \tag{64·5}$$
$$\Delta J = 0, \pm 1, \tag{64·6}$$

for problem *C*.

These rules are most readily deduced by algebraic methods like those used in Sec. 61*b* for the determination of the angular-momentum matrices.[1] Let $\vec{\mathcal{L}}_k$ and $\vec{D}_k$ denote respectively the vector operators for the orbital angular momentum and the electric moment of the *k*th electron. As a starting point we work out the Poisson brackets for the

[1] The first quantum-mechanical proof of these selection principles was the matrix derivation by M. Born, W. Heisenberg, and P. Jordan, *Zeits. f. Physik* **35**, 557 (1925).

three components of $\vec{\mathcal{L}}_k$ paired in turn with the three components of $\vec{D}_k$. These can be obtained by direct partial differentiation, or by application of the commutation rules (37·10) for the Cartesian coordinates and conjugate components of linear momentum.

$$
\begin{aligned}
[\mathcal{L}_{kx},D_{kx}] &= 0, & [\mathcal{L}_{ky},D_{kx}] &= -D_{kz}, & [\mathcal{L}_{kz},D_{kx}] &= D_{ky}, \\
[\mathcal{L}_{kx},D_{ky}] &= D_{kz}, & [\mathcal{L}_{ky},D_{ky}] &= 0, & [\mathcal{L}_{kz},D_{ky}] &= -D_{kx}, \\
[\mathcal{L}_{kx},D_{kz}] &= -D_{ky}, & [\mathcal{L}_{ky},D_{kz}] &= D_{kx}, & [\mathcal{L}_{kz},D_{kz}] &= 0.
\end{aligned}
\tag{64·7}
$$

Since all the coordinates and momenta of any one electron commute with all the coordinates and momenta of any other, the equations are equally valid if we drop the subscript k throughout.

Let $\hat{D}$ and $\hat{D}^\dagger$ denote the adjoint complex components of the electric moment defined by

$$
\hat{D} = D_x + iD_y = \sum_{k=1}^{f} e_k(x_k + iy_k), \qquad \hat{D}^\dagger = D_x - iD_y = \sum_{k=1}^{f} e_k(x_k - iy_k),
\tag{64·8}
$$

as in Sec. 55b. From Eqs. (64·7), and (64·8) it follows that

$$
[\mathcal{L}_z,\hat{D}] = i\hat{D}, \qquad [\mathcal{L}_z,\hat{D}^\dagger] = -i\hat{D}^\dagger.
\tag{64·9}
$$

Each of the above operator equations implies a corresponding matrix equation. Thus

$$
[\mathcal{L}_z,\hat{\mathbf{D}}] = i\hat{\mathbf{D}}, \qquad [\mathcal{L}_z,\hat{\mathbf{D}}^\dagger] = -i\hat{\mathbf{D}}^\dagger.
\tag{64.10}
$$

We assume a matrix scheme based on simultaneous eigenfunctions of $\mathcal{L}_z$ and such additional independent commuting observables μ, as are needed to make a complete set. As in Sec. 61b, p. 514, we number the various sets of mutually compatible eigenvalues of the μ's and call the general ordinal number τ. The typical matrix elements of $\hat{D}$ and $\mathcal{L}_z$ take the respective forms $\hat{D}(\tau',M_L'; \tau'',M_L'')$ and $M_L'\dfrac{h}{2\pi}\delta_{M_L'M_L''}\delta_{\tau'\tau''}$. Equations (64·10) yield

$$
\hat{D}(\tau',M_L'; \tau'',M_L'')(M_L'' - M_L' - 1) = 0,
\tag{64·11}
$$
$$
\hat{D}^\dagger(\tau', M_L'; \tau'',M_L'')(M_L'' - M_L' + 1) = 0.
\tag{64·12}
$$

Similarly, the relation $[\mathcal{L}_z,D_z] = 0$ yields

$$
D_z(\tau',M_L'; \tau'',M_L'')(M_L'' - M_L') = 0.
\tag{64·13}
$$

From these three equations it follows that all the elements of all three matrices $\mathbf{D}_x$, $\mathbf{D}_y$, $\mathbf{D}_z$ vanish in the above scheme, with the exception of those for which

$$
\Delta M_L = M_L'' - M_L' = 0, \pm 1.
$$

As emphasized at the beginning of Sec. 54f, the basic formulas (54·8), (54·9), and (54·10) for the transition probabilities due to dipole radiation assume that the vector electric-moment matrix $\vec{D}$ is to be computed from a complete orthonormal system of eigenfunctions of the atomic Hamiltonian. If the atomic Hamiltonian commutes with $\mathcal{L}_z$, there is a complete orthonormal system of simultaneous eigenfunctions of H and $\mathcal{L}_z$. When the transition probabilities are computed from the matrix of $\vec{D}$ in this scheme, there are no contributions from pairs of states which do not conform to the selection rule (64·3).

Since all the spin operators commute with all components of $\vec{\mathcal{L}}$, Eqs. (64·7) (64·9), and (64·10) are equally valid if we substitute for $\vec{\mathcal{L}}$ the total angular momentum $\vec{\mathcal{L}} + \vec{S}$, or $\vec{\mathcal{J}}$. Hence the matrices D_x, D_y, D_z in any scheme which makes J_z diagonal will have no nonvanishing elements which do not obey the selection rule (64·5).

There are two types of application of the selection rules for M_L and M, $viz.$, applications to free atomic systems with spherical symmetry, and applications to atomic systems with axial symmetry only. In the former case the selection rule for M_L in problem B and the selection rule for M in problem C (or B) are of use in working out the transition probabilities and intensities of different spectrum lines, but are not directly reflected in the elimination of corresponding lines because $\mathcal{L}_z$, or $\mathcal{J}_z$, as the case may be, is not a function of the energy, and pairs of states with different values of ΔM_L, or ΔM, contribute to the same spectrum line.

Consider, next, applications to atomic systems with axial symmetry only. The fixed-nuclei problem for a diatomic molecule (cf. Sec. 52a) and the Zeeman and Stark effects afford examples of this type of system in which the z component of angular momentum, say $\mathcal{L}_z$, becomes an integral of the motion when the z axis is identified with the axis of symmetry, although the perpendicular components are not integrals. In such a case every integral of the motion commutes with $\mathcal{L}_z^2$ (R_x and R_y, which do not commute with $\mathcal{L}_z$, may be integrals) and $\mathcal{L}_z^2$ becomes a function of the energy. We can accordingly label the energy levels with corresponding values of M_L^2, or $|M_L|$. The selection rule (64·3) is then directly reflected in the spectrum through the elimination of lines corresponding to changes in $|M_L|$ which have absolute values greater than unity.

In order to derive the selection rules (64·4) and (64·6) we shall employ a procedure first used by Dirac[1] in proving the selection rule for the two-particle problem of Sec. 55b and later slightly modified by Güttinger and Pauli[2] to cover the general case of any atomic system with spherical

[1] P. A. M. DIRAC, *Proc. Roy. Soc.* **A111**, 281 (1926); *P.Q.M.*, §47.

[2] P. GÜTTINGER and W. PAULI, *Zeits. f. Physik* **67**, 743 (1931); CONDON and SHORTLEY, *T. A. S.* **8³**, p. 60.

symmetry. From Eqs. (64·7) it follows that

$$[\mathcal{L}^2, D_z] = [\mathcal{L}_x{}^2, D_z] + [\mathcal{L}_y{}^2, D_z] = \mathcal{L}_y D_x + D_x \mathcal{L}_y - (\mathcal{L}_x D_y + D_y \mathcal{L}_x)$$

$$= 2\left\{(\mathcal{L}_y D_x - \mathcal{L}_x D_y) - \frac{h}{2\pi i} D_z\right\}$$

$$= 2(\mathcal{L}_y D_x - D_y \mathcal{L}_x) = 2(D_x \mathcal{L}_y - \mathcal{L}_x D_y).$$

By cyclic advancement of the subscripts:

$$[\mathcal{L}^2, D_x] = 2(D_y \mathcal{L}_z - \mathcal{L}_y D_z), \qquad [\mathcal{L}^2, D_y] = 2(D_z \mathcal{L}_x - \mathcal{L}_z D_x).$$

Hence

$$[\mathcal{L}^2, [\mathcal{L}^2, D_z]] = 2\left[\mathcal{L}^2, \mathcal{L}_y D_x - \mathcal{L}_x D_y - \frac{h}{2\pi i} D_z\right]$$

$$= 2\mathcal{L}_y[\mathcal{L}^2, D_x] - 2\mathcal{L}_x[\mathcal{L}^2, D_y] - 2\left(\frac{h}{2\pi i}\right)[\mathcal{L}^2, D_z]$$

$$= 4\mathcal{L}_y(D_y \mathcal{L}_z - \mathcal{L}_y D_z) - 4\mathcal{L}_x(\mathcal{L}_x D_z - D_x \mathcal{L}_z) + 2(\mathcal{L}^2 D_z - D_z \mathcal{L}^2)$$

$$= 4(\vec{\mathcal{L}} \cdot \vec{D})\mathcal{L}_z - 2(\mathcal{L}^2 D_z + D_z \mathcal{L}^2).$$

Thus

$$[\mathcal{L}^2, [\mathcal{L}^2, \vec{D}]] = 4(\vec{\mathcal{L}} \cdot \vec{D})\vec{\mathcal{L}} - 2(\mathcal{L}^2 \vec{D} + \vec{D}\mathcal{L}^2). \tag{64·14}$$

By direct expansion of the left-hand member,

$$\mathcal{L}^4 \vec{D} - 2\mathcal{L}^2 \vec{D}\mathcal{L}^2 + \vec{D}\mathcal{L}^4 = -4\left(\frac{h}{2\pi}\right)^2 (\vec{\mathcal{L}} \cdot \vec{D})\vec{\mathcal{L}} + 2\left(\frac{h}{2\pi}\right)^2 (\mathcal{L}^2 \vec{D} + \vec{D}\mathcal{L}^2). \tag{64·15}$$

Güttinger and Pauli observe that $\vec{\mathcal{L}} \cdot \vec{D}$ commutes with each component of $\vec{\mathcal{L}}$ and therefore with $\mathcal{L}^2$. It follows that $(\vec{\mathcal{L}} \cdot \vec{D})\vec{\mathcal{L}}$ commutes with $\mathcal{L}^2$. Hence the matrix of this quantity in a scheme which makes L^2 diagonal will have no nonvanishing elements which are off-diagonal with respect to L. Let us form the matrix element $(\tau', L'; \tau'', L'')$ of (64·15) in such a scheme. We again employ τ to designate the ordinal number of a set of eigenvalues of additional independent commuting observables needed for a complete set. Let us agree that $L' \neq L''$ in order to eliminate the last term of the equation. Then

$$\left(\frac{h}{2\pi}\right)^4 [(L')^2(L'+1)^2 - 2L'(L'+1)L''(L''+1) + (L'')^2(L''+1)^2]$$

$$\times \vec{D}(\tau', L'; \tau'', L'') = 2\left(\frac{h}{2\pi}\right)^4 [L'(L'+1) + L''(L''+1)]\vec{D}(\tau', L'; \tau'', L'').$$

This equation is reducible[1] to the form

$$[(L' + L'' + 1)^2 - 1][(L' - L'')^2 - 1]D(\tau', L'; \tau'', L'') = 0. \tag{64·16}$$

[1] *Cf.* references in footnote 2, p. 545.

Since $L' \neq L''$ and neither L' nor L'' can be negative, the first factor cannot vanish. Consequently $\vec{D}(\tau',L'; \tau'',L'')$ must vanish unless $L' - L'' = \pm 1$. We conclude that in general $\vec{D}(\tau',L'; \tau'',L'')$ is zero unless $L' - L'' = 0, \pm 1$. It follows that the selection rule (64·4) must hold for any atomic system if $\mathcal{L}^2$ is an integral of the motion. Similarly (64·6) must hold if $\mathcal{J}^2$ is an integral of the motion.

In the case of the two-particle problem without electron spin the scalar product $\vec{\mathcal{L}} \cdot \vec{D}$ vanishes automatically and Eq. (64·16) holds even when L' and L'' are equal. Transitions for which $\Delta L = 0$ are accordingly ruled out when L' and L'' do not both vanish. A general rule to be derived in the next paragraph forbids transitions of the latter type. Consequently the selection rule for the angular quantum number in this case becomes

$$\Delta L = \Delta l = \pm 1,$$

as proved in Sec. 55b by another method.

In concluding this section we derive rules forbidding transitions between two states for each of which $\mathcal{L}^2$ (or $\mathcal{J}^2$) has the eigenvalue zero. These rules originate in the fact that D_x, D_y, D_z are eigenfunctions of $\mathcal{L}^2$ and $\mathcal{J}^2$ with the eigenvalue $2(h/2\pi)^2$, i.e., with the quantum number L (or J) set equal to unity. Thus

$$(\mathcal{L}_x{}^2 + \mathcal{L}_y{}^2 + \mathcal{L}_z{}^2)\sum_{k=1}^{f} ex_k = \sum_{k=1}^{f}(\mathcal{L}_{kx}{}^2 + \mathcal{L}_{ky}{}^2 + \mathcal{L}_{kz}{}^2)ex_k$$

$$= \frac{eh}{2\pi i}\sum_{k=1}^{f}(\mathcal{L}_{ky}z_k - \mathcal{L}_{kz}y_k) = (1+1)\left(\frac{h}{2\pi}\right)^2\sum_{k=1}^{f}ex_k.$$

The product of the wave functions of two states for each of which $\mathcal{L}^2$ has the eigenvalue zero is also an eigenfunction of $\mathcal{L}^2$ with the eigenvalue zero. This product is consequently orthogonal to each component of $\vec{D}$. This proves the $\mathcal{L}^2$ rule for dipole radiation. The $\mathcal{J}^2$ rule follows in the same way. In fact it is possible by an argument of the above type to prove that these rules hold for electric quadrupole radiation as well.

65. THE HELIUM ATOM AND "EXCHANGE ENERGY"

65a. Two-electron Atoms.—The work of the previous section completes the theoretical derivation of the empirical spectroscopic rules formulated in Secs. 56, 57, and 58a. We shall now turn our attention to the simplest example of a many-electron atomic system with Russell-Saunders coupling. This is the case of a nucleus of charge Ze with two "planetary" electrons, i.e., the problem of the neutral helium atom and

of the helium-like ions of larger atomic number, such as Li^+, Be^{++}, etc. It was an investigation of this problem by Heisenberg[1] which first revealed the electrostatic nature of the energy differences between terms of different multiplicity with the same angular momentum and coming from the same configuration.   The investigation also led to the discovery of the restriction of physically admissible wave functions by the exclusion of those which are not antisymmetric with respect to electron interchanges (wave-mechanical form of the Pauli principle).

For our present purpose it is convenient to use symbols for the individual-electron wave functions in the central-field approximation which indicate their form more explicitly than those adopted in Sec. 63b. Explicit indications of the quantum numbers m_l and m_s are needed. The quantum numbers n, l can be replaced by a single index integer τ.   The notation of Sec. 61a will be used for the spin functions of the individual electrons.   There are eight orthogonal central-field product wave functions for a given pair of sets of values for the orbital quantum numbers n, l, m_l.   Four of these are derivable from the other four by application of the interchange permutation P_{12}. The four primary product functions can be described by the symbols

$$\left.\begin{aligned}
\phi_1 &= (1|\tau,m_l)\alpha(1)(2|\tau',m_l')\alpha(2),\\
\phi_2 &= (1|\tau,m_l)\beta(1)(2|\tau',m_l')\beta(2),\\
\phi_3 &= (1|\tau,m_l)\alpha(1)(2|\tau',m_l')\beta(2),\\
\phi_4 &= (1|\tau,m_l)\beta(1)(2|\tau',m_l')\alpha(2).
\end{aligned}\right\} \tag{65.1}$$

$\widehat{G}$ transforms ϕ_1, ϕ_2, ϕ_3, ϕ_4 into the following four independent antisymmetric functions

$$\widehat{G}\phi_1 = \frac{1}{\sqrt{2}}[(1|\tau,m_l)\alpha(1)(2|\tau',m_l')\alpha(2) - (2|\tau,m_l)\alpha(2)(1|\tau',m_l')\alpha(1)],$$

$$\widehat{G}\phi_2 = \frac{1}{\sqrt{2}}[(1|\tau,m_l)\beta(1)(2|\tau',m_l')\beta(2) - (2|\tau,m_l)\beta(2)(1|\tau',m_l')\beta(1)],$$

$$\widehat{G}\phi_3 = \frac{1}{\sqrt{2}}[(1|\tau,m_l)\alpha(1)(2|\tau',m_l')\beta(2) - (2|\tau,m_l)\alpha(2)(1|\tau',m_l')\beta(1)],$$

$$\widehat{G}\phi_4 = \frac{1}{\sqrt{2}}[(1|\tau,m_l)\beta(1)(2|\tau',m_l')\alpha(2) - (2|\tau,m_l)\beta(2)(1|\tau',m_l')\alpha(1)].$$

Two of these functions can be resolved at once into products of orbital and spin functions.   Thus

$$\left.\begin{aligned}
\Phi_1 &= \widehat{G}\phi_1 = \frac{1}{\sqrt{2}}[(1|\tau,m_l)(2|\tau',m_l') - (2|\tau,m_l)(1|\tau',m_l')]\alpha(1)\alpha(2),\\
\Phi_2 &= \widehat{G}\phi_2 = \frac{1}{\sqrt{2}}[(1|\tau,m_l)(2|\tau',m_l') - (2|\tau,m_l)(1|\tau',m_l')]\beta(1)\beta(2).
\end{aligned}\right\} \tag{65.2}$$

[1] W. Heisenberg, *Zeits. f. Physik* **38**, 411 (1926), **39**, 499 (1926), **41**, 239 (1927).

By addition, subtraction, and renormalization one obtains two other antisymmetric functions which factor in the same way, *viz.*,

$$\Phi_3 = \frac{1}{\sqrt{2}}(\mathcal{G}\phi_3 + \mathcal{G}\phi_4) = \frac{1}{2}[(1|\tau,m_l)(2|\tau',m_l') - (2|\tau,m_l)(1|\tau',m_l')]$$
$$\times [\alpha(1)\beta(2) + \beta(1)\alpha(2)],$$
$$\Phi_4 = \frac{1}{\sqrt{2}}(\mathcal{G}\phi_3 - \mathcal{G}\phi_4) = \frac{1}{2}[(1|\tau,m_l)(2|\tau',m_l') + (2|\tau,m_l)(1|\tau',m_l')]$$
$$\times [\alpha(1)\beta(2) - \beta(1)\alpha(2)]. \tag{65·3}$$

Φ_4 is symmetric with respect to a permutation of the space coordinates but antisymmetric in the spin coordinates, while the other three functions are antisymmetric in the space coordinates and symmetric in the spin coordinates. Φ_4 is readily seen to be an eigenfunction of S^2 with the eigenvalue 0, while Φ_1, Φ_2, Φ_3 are eigenfunctions of S^2 with the common eigenvalue $2(h/2\pi)^2$. Thus Φ_4 belongs to the singlet system ($S = 0$), while the other three belong to the triplet system ($S = 1$). All four of the Φ's are eigenfunctions of $\mathcal{L}_z$ and S_z as well, with quantum numbers obtainable by inspection and indicated in Table IV.[1] Thus in this case the process of antisymmetrizing and factoring gives a set of wave functions which make S^2, S_z, L_z diagonal.

<div align="center">TABLE IV</div>

	Φ_1	Φ_2	Φ_3	Φ_4
S	1	1	1	0
M_S	1	-1	0	0
M_L	$m_l + m_l'$	$m_l + m_l'$	$m_l + m_l'$	$m_l + m_l'$

If possible, we should like L^2 to be diagonal also, since the order of the secular determinants to be solved would, then be reduced to a minimum. In fact it is easy to see that the order of the secular determinants would thereby be reduced to unity. We know from Sec. 63*d* that the coupling of *two* electrons produces only one term of a given type originating in a given initial configuration. Hence all simultaneous eigenfunctions of $\mathcal{L}^2$, S^2, $\mathcal{L}_z$, S_z coming from a common configuration and having the same set of values of L, S, M_L, M_S are linearly dependent. It follows that by dealing with one such set at a time we reduce the order of the secular determinant to unity, as was to be proved. In other

---

[1] From the equality of the space factors of Φ_1, Φ_2, Φ_3 we can at once verify the theorem that the energy is independent of M_S. The final zero-order wave functions will be made up of linear combinations of Φ's with the same S, M_S, M_L. The space factors of the Φ's with given S and M_L are independent of M_S. Hence the secular equation (U_S neglected) and the energies, which are its roots, are independent of M_S.

words the first-order approximation to the energy of a problem-B level
with given L,S values coming from a given configuration is simply the
mean value of the Hamiltonian $H_B = H_0 + H_1$ for the unique correspond-
ing case A wave function.

Although not all the Φ's are eigenfunctions of $\mathcal{L}^2$, there is a very
important class of Φ's which are, *viz.*, those which come from configura-
tions in which one electron is in an s state with azimuthal quantum
number zero.   Suppose, for example, that

$$(1|\tau,m_l) = (1|n,l,m_l) \quad = (1|n,0,0) \equiv v(1), \atop (2|\tau',m_{l'}) = (2|n',l',m_{l'}) \equiv w(2).} \qquad (65\cdot4)$$

Then

$$\mathcal{L}^2\phi_\lambda = [(\mathcal{L}_{1x} + \mathcal{L}_{2x})^2 + (\mathcal{L}_{1y} + \mathcal{L}_{2y})^2 + (\mathcal{L}_{1z} + \mathcal{L}_{2z})^2]v(1)w(2)$$

$$\times \text{ spin factor } = l'(l' + 1)\left(\frac{h}{2\pi}\right)^2 \phi_\lambda. \qquad \lambda = 1, 2, 3, 4.$$

Φ_λ is also seen to be an eigenfunction of $\mathcal{L}^2$ with the eigenvalue l'.   The
normal state of helium and the helium-like ions, together with all states
in which only one electron is excited, belongs to the class under con-
sideration.   Since all the more familiar optical energy levels are of this
type, we can exclude other cases from consideration without great
practical loss of generality.

Let $E^{(0)}$ and $E_\lambda^{(1)}$ denote respectively the central-field energy of a
configuration and the first-order energy correction for an associated Φ_λ.
It follows from the above that their sum, say $E_\lambda^{(0+1)}$, is given by

$$E_\lambda^{(0+1)} = (H_B\Phi_\lambda,\Phi_\lambda) = ([H_0 + H_1]\Phi_\lambda,\Phi_\lambda). \qquad (65\cdot5)$$

Summing over the different pairs of values of the spin variables we reduce
the scalar product to the form

$$E_\lambda^{(0+1)} = \int \cdots \int_\infty [v(1)w(2) \pm w(1)v(2)]^*H_B[v(1)w(2)$$

$$\pm w(1)v(2)]d\tau_1 d\tau_2, \qquad (65\cdot6)$$

where the negative signs go with the values 1, 2, 3 for λ and the positive
signs go with the value 4.   Subtracting off the central-field energy, we
have

$$E_\lambda^{(1)} = I_1 \pm I_2. \qquad (65\cdot7)$$

$$I_1 = 2\int \cdots \int H_1|v(1)|^2|w(2)|^2 d\tau_1 d\tau_2. \qquad (65\cdot8)$$

$$I_2 = 2\int \cdots \int w(1)^*v(2)^*H_1v(1)w(2)d\tau_1 d\tau_2. \qquad (65\cdot9)$$

In first approximation I_1 gives the difference between the central-field
energy and the mean of the singlet and triplet energies. I_2 determines

the spacing of the corresponding singlet and triplet levels. It is the analogue for the helium problem of the integral I_2 met with in the study of the H_2 molecule problem in Sec. 52b and is accordingly referred to as an exchange integral.

In the special case under consideration (62·10) degenerates into

$$H_1 = \frac{e^2}{r_{12}} - \left[\frac{Ze^2}{r_1} + \frac{Ze^2}{r_2} + V(r_1) + V(r_2) \right].$$

We have to decide what form to give $V(r)$ before we can proceed farther. Ordinarily in an atomic computation it is necessary to apply the Hartree self-consistent field method in order to determine a satisfactory central-field potential function V. If we exclude states for which l' is zero, however, we can adopt a simple suggestion due to Heisenberg which avoids this complication. The suggestion is, in effect, to make two choices of $V(r)$, viz., $V_1(r) = -\dfrac{Ze^2}{r}$, and $V_2(r) = -\dfrac{(Z-1)e^2}{r}$, using one for the inner electron and the other for the outer electron. The basis for this procedure is given by the small values of the quantum defect for the principal-series energy levels (0.06 for the triplet levels and -0.01 for the singlet levels) indicating a small interpenetration of the wave functions of the inner electron and the outer electron when the latter has an azimuthal quantum number not less than unity. Neglecting interpenetration entirely, the inner electron would be subject, on the average, to the field of the bare nucleus, while the outer electron would be subject, on the average, to the field obtained by condensing the inner one upon the nucleus. Consequently a good product-form approximate solution of the Hamiltonian H_B should be $v_Z(1)w_{Z-1}(2)$, where v_Z is a $1s$ function for the bare nucleus, and w_{Z-1} is an n', l', m_l' function for the nucleus of atomic number $Z - 1$. Using these initial approximations we can form functions Φ_1, Φ_2, Φ_3, Φ_4 by permutation and linear combination, as before, but they are not then eigenfunctions of any H_0. In fact $v_Z(1)w_{Z-1}(2)$ is an eigenfunction of

$$H_{01} = -\frac{h^2}{8\pi^2\mu}(\nabla_1{}^2 + \nabla_2{}^2) - \frac{Ze^2}{r_1} - \frac{(Z-1)e^2}{r_2}, \qquad (65\cdot10)$$

while the permuted function $v_Z(2)w_{Z-1}(1)$ is an eigenfunction of

$$H_{02} = -\frac{h^2}{8\pi^2\mu}(\nabla_1{}^2 + \nabla_2{}^2) - \frac{(Z-1)e^2}{r_1} - \frac{Ze^2}{r_2}. \qquad (65\cdot11)$$

A linear combination is an eigenfunction of neither H_{01} nor H_{02}. This fact takes the calculation out of the domain of the conventional perturbation theory of Sec. 48 and puts it into the domain of the variational method of Sec. 51.

The procedure is the same as for the Heitler-London calculation of the energy of the H_2 molecule (Sec. 52b). Equation (65·6) is still applicable, but its reduction takes a slightly different form. We have

$$H_B v(1)w(2) = H_{01}v(1)w(2) + H_{11}v(1)w(2) = [E^{(0)} + H_{11}]v(1)w(2),$$
$$(65\cdot12)$$

$$H_B v(2)w(1) = [E^{(0)} + H_{12}]v(2)w(1), \qquad (65\cdot13)$$

where

$$H_{11} = \frac{e^2}{r_{12}} - \frac{e^2}{r_2}, \qquad H_{12} = \frac{e^2}{r_{12}} - \frac{e^2}{r_1}. \qquad (65\cdot14)$$

The reduction of (65·6) with the above yields

$$E_\lambda^{(1)} = \int \cdots \int_\infty H_{11}|v(1)|^2|w(2)|^2 d\tau_1 d\tau_2 +$$

$$\int \cdots \int_\infty H_{12}|v(2)|^2|w(1)|^2 d\tau_1 d\tau_2 \pm \int \cdots \int_\infty H_{11}v(1)v(2)^*w(1)w(2)^* d\tau_1 d\tau_2$$

$$\pm \int \cdots \int_\infty H_{12}v(2)v(1)^*w(2)w(1)^* d\tau_1 d\tau_2.$$

The first integral is equal to the second and the third to the fourth. Let

$$I_1' = 2\int \cdots \int H_{11}|v(1)|^2|w(2)|^2 d\tau_1 d\tau_2; \quad \left.\right\}$$
$$I_2' = 2\int \cdots \int H_{11}v(1)v(2)^*w(1)w(2)^* d\tau_1 d\tau_2. \quad \left.\right\} \qquad (65\cdot15)$$

Then

$$E_\lambda^{(1)} = I_1' - I_2' \qquad (\lambda = 1,\ 2,\ 3); \qquad E_4^{(1)} = I_1' + I_2'. \qquad (65\cdot16)$$

The first-order approximate energy determination is thus reduced to the problem of working out two sextuple integrals. The reader is referred to Heisenberg's paper[1] for the details of the integration. The most important result of this work is the conclusion that the exchange integral I_2' is positive. It follows that singlet-system energy levels should lie *above* the corresponding triplet levels in accordance with experiment. The calculated differences between corresponding singlet and triplet energies are in fair agreement with experiment considering the roughness of the computation. Thus the computed difference between the quantum defects of the term $1s\ 2p\ ^3P$ and the term $1s\ 2p\ ^1P$ checks with the observed difference to within 20 per cent. The agreement for the $1s\ 3p\ ^3P$ and $1s\ 3p\ ^1P$ terms is within 15 per cent, while that for the corresponding terms of the Li^+ ion is within 10 per cent.

The calculation of the absolute positions of the terms is not at all good but becomes reasonably satisfactory when the first-order corrections are supplemented by those of second order.

The helium-atom problem is fundamentally of the same type as the fixed-nuclei hydrogen-molecule problem, since in each case we have to do with two electrons in a fixed external force field. In dealing with H_2

[1] *Cf.* second reference, footnote 1, p. 548.

in Sec. 52 we used spin-free wave functions. This procedure is legitimate when there are not more than two electrons, and the spin-orbit energy U_s is omitted from the Hamiltonian, because the orbital wave functions derived from the spin-free theory can be antisymmetrized by the introduction of properly chosen spin factors. Thus, since the orbital wave function ψ' of (52·4) is symmetric with respect to an interchange of the space coordinates, it suffices to add the antisymmetric spin factor $\alpha(1)\beta(2) - \beta(1)\alpha(2)$ in order to satisfy the Pauli principle. Similarly the antisymmetric orbital function ψ'' can be correlated with any one of the three symmetric spin factors $\alpha(1)\alpha(2)$, $\beta(1)\beta(2)$, $\alpha(1)\beta(2) + \beta(1)\alpha(2)$ to satisfy the Pauli principle. Hence the energy level E' is of the singlet-system type, while E'' belongs to the triplet type. In this case the order of the levels is opposite to that for helium, the singlet level being lower, because the exchange integral has the opposite sign.

65b. The Exchange Phenomenon.—The "cause" of the energy difference of corresponding singlet and triplet terms in both of these two-electron problems is to be found in (a) the permutation degeneracy of the zero-order product-form wave functions and in (b) the removal of the permutation operator from the list of integrals of the motion when the e^2/r_{12} term in the potential energy is introduced. From the wave point of view we have to do with the splitting of a degenerate vibration frequency when the symmetry which produced the degeneracy is removed. This is entirely analogous to the frequency splitting which accompanies the coupling of two resonating electrical circuits. In the electrical case the coupling results in a secular surging of the energy from one circuit into the other and back again when the coupled system is set into oscillation by the excitation of one circuit only. The frequency of these surge oscillations is equal to the difference between the two possible normal vibration frequencies of the coupled system. Essentially the same phenomenon is possible in these atomic problems.

Let ψ_1 and ψ_2 be eigenfunctions of the Hamiltonian of a two-electron atomic system belonging to neighboring energy levels E_1 and E_2, respectively. Let the system be started off in a subjective state with the wave function $(\psi_1 + \psi_2)/\sqrt{2}$. The corresponding solution of the second Schrödinger equation is

$$\Psi = \frac{1}{\sqrt{2}}\left[\psi_1 e^{-\frac{2\pi i}{h}E_1 t} + \psi_2 e^{-\frac{2\pi i}{h}E_2 t}\right].$$

A simple reduction gives

$$\Psi = \frac{1}{\sqrt{2}}e^{-\frac{2\pi i}{h}\left(\frac{E_1+E_2}{2}\right)t}\left[(\psi_1 + \psi_2)\cos\frac{2\pi}{h}\left(\frac{E_1-E_2}{2}\right)t \right.$$
$$\left. + (\psi_2 - \psi_1)\sin\frac{2\pi}{h}\left(\frac{E_1-E_2}{2}\right)t\right]. \quad (65·17)$$

It follows that at times which are *even* multiples of $h/2(E_1 - E_2)$ the probability density $|\Psi|^2$ reduces to $|\psi_1 + \psi_2|^2/2$, while at times which are *odd* multiples of $h/2(E_1 - E_2)$ the probability density reduces to $|\psi_1 - \psi_2|^2/2$. Thus the system oscillates back and forth with the beat frequency $(E_1 - E_2)/h$ from the state $(\psi_1 + \psi_2)/\sqrt{2}$ to the state $(\psi_1 - \psi_2)/\sqrt{2}$.

In the problem B approximation these beats lead to a periodic exchange of roles on the part of the two electrons *provided* that we arbitrarily omit the spin coordinates from the wave functions and treat the problem as we treated the fixed-nuclei H_2 molecule problem in Sec. 52. (The spin-free wave functions permit the electrons to have different roles because they are not necessarily either symmetric or antisymmetric with respect to an interchange of coordinates, although they are eigenfunctions of H_B with the eigenvalues allowed by the Pauli principle.)

Let us assume, for example, that ψ_1 is the product of a symmetric function u_1 of the spatial coordinates and an antisymmetric function of the spin coordinates, whereas ψ_2 is the product of an antisymmetric space function u_2 and a symmetric spin function. Dropping the spin factors, we assume the initial wave function $(u_1 + u_2)/\sqrt{2}$ and apply (65·17). At times which are even multiples of $h/2(E_1 - E_2)$ the probability density $|\Psi|^2$ takes the form $|u_1 + u_2|^2/2$, while at times which are *odd* multiples of the same unit, $|\Psi|^2$ takes the form $|u_1 - u_2|^2/2$. But the permutation operator P_{12} converts $(u_1 + u_2)/\sqrt{2}$ into $(u_1 - u_2)/\sqrt{2}$ and $(u_1 - u_2)/\sqrt{2}$ into $(u_1 + u_2)/\sqrt{2}$. *Thus the passage of the time interval $h/2(E_1 - E_2)$ is equivalent in effect to the application of the permutation P_{12} and can be said to interchange the roles of the two electrons in the wave function.* This exchange of places becomes more concrete and vivid if we make use of the case A approximations

$$\frac{1}{\sqrt{2}}[v(1)w(2) + w(1)v(2)] \quad \text{and} \quad \frac{1}{\sqrt{2}}[v(1)w(2) - w(1)v(2)]$$

of Eq. (65·6) for u_1 and u_2, respectively. With the identification,

$$\frac{1}{\sqrt{2}}(u_1 + u_2) = v(1)w(2), \qquad \frac{1}{\sqrt{2}}(u_1 - u_2) = w(1)v(2).$$

Thus the resonance phenomenon results in a periodic exchange of the identities of the electrons associated with the central-field states indicated by v and w.

In the case of the H_2 molecule problem of Sec. 52b we can identify u_1 and u_2 in first approximation with the functions ψ' and ψ'' of Eq. (52·4). Equation (65·17) then leads to a periodic exchange in the identities of the electrons associated with the two nuclei. It should be clearly understood, however, that these exchange "pictures" cannot be formulated in

terms of the complete antisymmetric wave functions with spin argu-
ments included.   Moreover, it is certainly wrong to think of the periodic
exchange of places of the electrons as the *cause* of the resonance-energy
difference $E_1 - E_2$.

66. DIAGONAL SUMS AND THE PROBLEM B ENERGY LEVELS

The procedure used in Sec. 65 for securing zero-order (central-field)
antisymmetric wave functions which diagonalize L^2 and S^2 is a very
special one applicable only to the particular problem there considered.
(It is impossible to set up complete antisymmetric wave functions which
are the products of space and spin factors if there are more than two
electrons in the system.   There are only two kinds of individual-electron
spin functions and hence it is not possible to make an antisymmetric
spin function involving more than two electrons.)

· There are a number of more general procedures, however, for locating
the energy levels of problem B in first-order approximation.   The first
and most obvious of these is the following.   Let the Slater wave functions
$\Phi_\alpha = \hat{\mathsf{G}}\phi_\alpha$ for the configuration under consideration be listed.   For a
given pair of values of M_L and M_S there will be a finite number of func-
tions,   say $\Phi_1{}^{(M_L M_S)}$, $\Phi_2{}^{(M_L M_S)}$, . . . , $\Phi_\lambda{}^{(M_L M_S)}$.   The   finite   energy
matrix $^{(M_L M_S)}\mathbf{H}_1$, based on these functions, can then be worked out and
diagonalized by solving the corresponding secular equation (48·7).

In practice the calculation is greatly simplified by an observation
due to Slater (*loc. cit.*, footnote 3, p. 535).   We introduce the modified
procedure by the consideration of the special case where there are two
equivalent p electrons outside the closed shells.   Table III, p. 540, shows
the possible M_L and M_S values with the designations of the Slater wave
functions for each.

Suppose we begin with the square of the table for which $M_L = 2$,
$M_S = 0$.   There is a *single* Slater wave function, which must accordingly
be a simultaneous eigenfunction of $\mathcal{L}^2$ and S^2.   It belongs to the ^1D term
as previously noted.   The corresponding secular determinant is of the
first order and the energy is given by the diagonal element

$$^{(2,0)}H_1(1,1) = (H_1\Phi_1{}^{(2,0)},\Phi_1{}^{(2,0)}).$$

Similarly the columns of the table for $M_L = 1$ and $M_L = -1$ have one
wave function in each square.   The mean value of H_1 for any of these
functions gives the energy of the ^3P state.   When this is done there
remains only the ^1S state to be calculated.   The first-order energy
correction for this state is one of the roots of the secular determinant
for the three Slater wave functions

$$\Phi_1{}^{(0,0)} = (1^+\,0^-), \qquad \Phi_2{}^{(0,0)} = (0^+\,\dot{0}^-), \qquad \Phi_3{}^{(0,0)} = (1^-\,-1^+),$$

the other two roots being the previously computed first-order energy

corrections of the ^1D and ^3P states. Fortunately it is unnecessary to compute the off-diagonal elements of the matrix $^{(0,0)}\mathbf{H}_1$ for this square of the table, since the sum of the diagonal terms of a finite square matrix is invariant of a canonical transformation (*cf.* Sec. 44*b*). Hence the sum of the diagonal elements of $^{(0,0)}\mathbf{H}_1$ is equal to the sum of the roots of the equation. [From elementary algebra we know that the sum of the roots of an algebraic equation of the form

$$x^n + A_1 x^{n-1} + A_2 x^{n-2} + \cdots + A_n = 0$$

is $-A_1$. The expansion of the secular determinant shows that this coefficient is $\sum_{k=1}^{3} {}^{(0,0)}H_1(k,k)$.] Therefore the energy correction for the ^1S state is obtainable by subtracting off the values of $E^{(1)}$ for the ^1D and ^3P states from the diagonal sum in question. Thus the location of the energy levels for the special case under consideration is reduced in first approximation to the calculation of the diagonal terms of the energy matrix based on Slater wave functions.

In most cases the application of the procedure sketched above suffices to eliminate the necessity of computing any off-diagonal matrix elements. However, if there are two or more terms of the same type originating in the same configuration, it is necessary to work out some off-diagonal elements. In this case every secular equation involving one of the terms will involve the other as well. Consequently it becomes necessary either to evaluate the off-diagonal matrix elements of $\mathbf{H}_1$ needed for the solution of one of these secular equations, or to solve the simpler secular equation which takes its place when the Slater wave functions are replaced by others which diagonalize L^2 and S^2.

For more complete and detailed information regarding the theory of atomic spectra the reader is again referred to the comprehensive work on that subject by Condon and Shortley.

APPENDIX A

THE CALCULUS OF VARIATIONS AND THE PRINCIPLE OF LEAST ACTION[1]

The Fundamental Problem.—The calculus of variations deals with a generalization of the ordinary problem of maxima and minima of a function of one or more independent variables. This generalization permits the reformulation of problems involving differential equations as problems in maxima and minima.

The fundamental problem of the calculus of variations in its simplest form is the following. Let $F(x,y,t)$ denote a continuous and twice differentiable function of the three arguments x, y, t. Let x be a continuous differentiable function of t, and let y be identified with the derivative of x with respect to t which we denote by $\dot{x}$. Then $F(x,\dot{x},t)$ becomes an implicit function of the single argument t, although explicitly a function of the three variables x, $\dot{x}$, t. Consider the integral

$$J[x] = \int_{t_0}^{t_1} F(x,\dot{x},t)dt. \tag{A-1}$$

J depends on the form of the function $x(t)$ but is not a function of x in the ordinary sense since there is no one-to-one correlation of values of x with values of J. We call J a function of a function, or a functional. It is required to find those functions $x(t)$ having continuous second derivatives which yield a maximum or a minimum value of J when we consider only comparison functions with the same terminal values $x(t_0) = x_0$, $x(t_1) = x_1$.

Euler's Equation.—In the case of an ordinary function y of an independent variable x, a *necessary* condition for a maximum or minimum is that dy/dx or the differential dy shall vanish at the point under consideration. A similar necessary condition for the minimization of J can be set up and yields a differential equation (Euler's) for $x(t)$. A solution of this equation subject to the given terminal conditions may yield either a maximum or minimum value of J or a generalized "point of
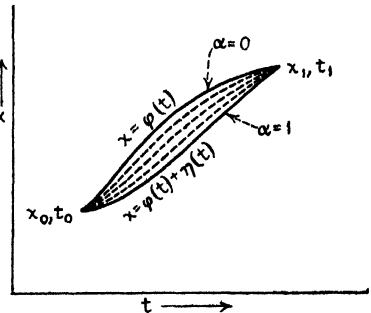
Fig. 26.

---

[1] *Cf.* Sec. 3, p. 9.

inflection." For many physical applications it makes no difference which.

In order to formulate Euler's equation we form the one-parameter family of comparison functions

$$x(t,\alpha) = \varphi(t) + \alpha\eta(t), \tag{A-2}$$

in which $\varphi(t)$ satisfies the terminal conditions and $\eta(t)$ is any continuous and twice-differentiable function which vanishes at t_0 and t_1. Inserting $x(t,\alpha)$ for x in Eq. (A-1) we convert J into a function of α.

$$J(\alpha) = \int_{t_0}^{t_1} F\left(x(t,\alpha), \frac{\partial}{\partial t}x(t,\alpha), t\right)dt. \tag{A-3}$$

Let us now assume that the function $x = \varphi(t)$ yields either a maximum or a minimum value of J. It follows that $\partial J/\partial\alpha$ vanishes when α is zero, or, what amounts to the same thing, that the *principal part* of the increment of J due to an arbitrary increment $\delta\alpha$ in α is zero if the initial value of α is zero. This principal part is indicated by the symbol δJ and is called the first variation of J. Its value is by definition

$$\delta J = \left(\frac{dJ}{d\alpha}\right)_{\alpha=0} \delta\alpha. \tag{A-4}$$

Similarly if Q is any quantity depending on the function $x(t)$ we define the *first variation of Q* by the equation

$$\delta Q = \left(\frac{\partial}{\partial\alpha}Q[x(t,\alpha)]\right)_{\alpha=0} \delta\alpha.$$

It will be observed that since $x(t,\alpha)$ involves $\varphi(t)$ and $\eta(t)$, δQ depends on the choice of these two functions. To avoid confusion regarding this choice we shall sometimes replace the symbol δQ by the more explicit symbol $\delta_\varphi{}^\eta Q$.

Applying the above general definition of δQ to the quantities x and x' we obtain

$$\delta x = \left[\frac{\partial}{\partial\alpha}x(t,\alpha)\right]_{\alpha=0}\delta\alpha = \eta(t)\delta\alpha,$$

$$\delta\dot{x} = \left[\frac{\partial}{\partial\alpha}\dot{x}(t,\alpha)\right]_{\alpha=0}\delta\alpha = \frac{d}{dt}\delta x.$$

In these special cases δQ is independent of $\varphi(t)$, *i.e.*, of the unvaried expression for $x(t)$.

It is easy to prove that all the ordinary rules of the differential calculus apply to the computation of variations. Hence

$$\delta F = \frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial\dot{x}}\delta\dot{x} = \frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial\dot{x}}\frac{d}{dt}\delta x. \tag{A-5}$$

The limits of integration for J are independent of α and we may differentiate under the integral sign to obtain

$$\delta J = \int_{t_0}^{t_1}\left[\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial \dot{x}}\frac{d}{dt}\delta x\right]dt. \tag{A-6}$$

On integration by parts this yields

$$\delta J = \int_{t_0}^{t_1}\left[\frac{\partial F}{\partial x} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right)\right]\delta x\,dt + \frac{\partial F}{\partial \dot{x}}\delta x\Big|_{t_0}^{t_1}$$

$$= \int_{t_0}^{t_1}\left[\frac{\partial F}{\partial x} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right)\right]\delta x\,dt. \tag{A-7}$$

In equations (A-5), (A-6), and (A-7) the partial derivatives of F are to be carried out treating F as a function of the three independent variables x, $\dot{x}$, t, while the symbol d/dt denotes a differentiation with respect to t treated as the sole independent variable.

If $x = \varphi(t)$ gives a true maximum or minimum for J it is clear that $\delta_\varphi{}^\eta J$ must vanish for every choice of the *arbitrary* function $\delta x = \alpha\eta(t)$. J is then said to be *stationary*. Then the cofactor of δx in the right-hand member of Eq. (A-7) must vanish identically. In other words, $x = \varphi(t)$ must be a solution of

$$\frac{\partial F}{\partial x} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right) = 0. \tag{A-8}$$

This is Euler's equation.

The converse proposition that solutions of Eq. (A-8) which satisfy the terminal conditions yield either a maximum or minimum value of J is not correct, but we can say that such solutions give zero for the first variation in J. They are called *extremals* and are of great intrinsic importance in mathematical physics independent of the fact that they include all functions which give maximum or minimum J values.

Example I.—Let T and V denote respectively the kinetic and potential energies of a particle moving in one dimension along a straight line. Let L denote the Lagrangian function or kinetic potential $T - V$. Identifying F with L we see that the process of "extremalizing" the integral

$$J = \int_{t_0}^{t_1}[\tfrac{1}{2}\mu\dot{x}^2 - V(x)]dt$$

is equivalent to solving the differential equation

$$-\frac{\partial}{\partial x}\left[\frac{1}{2}\mu\dot{x}^2 - V\right] + \frac{d}{dt}\left\{\frac{\partial}{\partial \dot{x}}\left[\frac{1}{2}\mu\dot{x}^2 - V\right]\right\} = \mu\ddot{x} + \frac{\partial V}{\partial x} = 0.$$

Thus Newton's law of motion is identical with Euler's equation for the integral J. This is a special case of Hamilton's principle.

Several Dependent Variables.—A somewhat more general problem is that of "extremalizing" the integral J, when F has a number of argu-

ment functions. For example, we may assume that F depends on x_1, $\dot{x}_1$, x_2, $\cdots$ x_n, $\dot{x}_n$, t. Then a slight extension of the foregoing argument shows that the functions define an extremal trajectory or curve when they satisfy the set of simultaneous equations

$$\frac{\partial F}{\partial x_i} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}_i}\right) = 0. \qquad i = 1, 2, \cdots n \qquad (A-9)$$

If F is identified with the Lagrangian function or kinetic potential L in a multidimensional mechanical problem, the above Euler equations become the Lagrangian equations of classical mechanics. We conclude that if the equations of motion can be reduced to the Lagrangian form (A-9) by the introduction of a suitably defined function L having the coordinates, their velocities, and the time as arguments, the trajectory of the system executing its natural motion is such that

$$\delta \int_{t_0}^{t_1} L(q,\dot{q},t)dt = 0. \qquad (A-10)$$

This is Hamilton's principle in one of its more general forms. The comparison trajectories allowed in the formulation of this principle are restricted to those which carry the system from a common initial configuration A to a common final configuration B in the common time interval $t_1 - t_0$. If the forces acting on the system are derivable from a function V of the coordinates and the time, Lagrange's equations and Hamilton's principle are always applicable with

$$L = T - V.$$

Principle of Least Action.—The principle of least action in its usual form is restricted to conservative systems, *i.e.*, to systems subject to forces derivable from a potential function V which is independent of the time. To set up this principle we consider the time integral of a multiple of the kinetic energy T of the system over its trajectory from an initial configuration A to a final configuration B. The value of the integral for the natural trajectory is stationary with respect to values obtained from neighboring non-mechanical trajectories having the same total energy and carrying the system from the same initial configuration to the same final configuration, but *not* necessarily in the same time.

We shall verify the principle for the case of a single particle by showing that it is equivalent to Newton's law of motion. Let E be the total energy and let ds denote an element of arc. The action integral S for the path AB is defined as follows:

$$S = \int_A^B 2T dt = \int_A^B \mu v ds = \int_A^B p ds. \qquad (A-11)$$

Here p denotes the absolute value of the momentum for the point x,y,z,

determined by the energy equation

$$\frac{p^2}{2\mu} = E - V(x,y,z).$$

The principle of least action requires that $\delta S = 0$ for the natural orbit when the trajectory is varied, but E is held constant.

Usually this makes S an actual minimum, but there are exceptions.[1] We cannot apply Euler's equations directly to S as given by Eq. (A-11) since it is not in the standard form of Eq. (A-1). However, it may be carried over into the standard form by the introduction of an auxiliary variable u which increases monotonically from 0 to 1 as the particle moves from A to B. Let

$$x = f_1(u), \qquad y = f_2(u), \qquad z = f_3(u),$$
$$ds = \left[\left(\frac{dx}{du}\right)^2 + \left(\frac{dy}{du}\right)^2 + \left(\frac{dz}{du}\right)^2 \right]^{\frac{1}{2}} du,$$
$$p = \sqrt{2\mu[E - V(x,y,z)]}.$$

Denoting differentiation with respect to u by a prime, we have

$$S = \int_0^1 p(E,x,y,z)[(x')^2 + (y')^2 + (z')^2]^{\frac{1}{2}} du. \qquad \text{(A-12)}$$

As this expression for S is in the standard form we may write down Euler's equation for the x coordinate:

$$\frac{\partial}{\partial x}\{p[(x')^2 + (y')^2 + (z')^2]^{\frac{1}{2}}\} - \frac{d}{du}\left(\frac{\partial}{\partial x'}\{p[(x')^2 + (y')^2 + (z')^2]^{\frac{1}{2}}\}\right) = 0.$$

On reduction we obtain

$$\frac{\partial p}{\partial x} - \frac{d}{ds}\left(p\frac{dx}{ds}\right) = 0. \qquad \text{(A-13)}$$

Multiply through by p/μ and make use of the relations

$$\frac{\mu ds}{p} = \frac{ds}{v} = dt, \qquad \left(\frac{p}{\mu}\right)\frac{\partial p}{\partial x} = -\frac{\partial V}{\partial x}.$$

Equation (A-13) takes the final form

$$\frac{\partial V}{\partial x} + \mu\frac{d^2x}{dt^2} = 0,$$

which brings us back to Newton's law. As the other three equations reduce in the same way, we conclude that the principle of least action is equivalent to Newton's second law of motion.

[1] *Cf.* E. T. WHITTAKER, *Analytical Dynamics*, §103, p. 250, Cambridge, 1917.

The distinguishing feature of the principle of least action is that it leads to purely geometric (time-free) orbital differential equations (A-13). Both Hamilton's principle and the principle of least action owe their importance primarily to the fact that they afford formulations of the dynamical laws which are independent of any reference to particular coordinate systems.

Least Action with Variable Mass.—In order to extend the principle of least action to problems in which the mass of the particle varies with its speed according to the relativistic law

$$\mu = \frac{\mu_0}{\sqrt{1 - \dfrac{v^2}{c^2}}}, \tag{A-14}$$

we define the action integral by the equation [*cf.* (A-11)]

$$S = \int_A^B \mu v\, ds = \int_A^B p\, ds, \tag{A-15}$$

although in this case $\mu v ds$ is not equivalent to $2T dt$. To express the momentum p in terms of the space coordinates the Newtonian energy equation must be replaced by the relativistic equation from which we obtain

$$p(x,y,z,E_r)^2 = \frac{1}{c^2}\{[E_r - V(x,y,z)]^2 - \mu_0^2 c^2\}. \tag{A-16}$$

As in the Newtonian case, we derive Eq. (A-13) from the definition of S and this reduces immediately to the form

$$\frac{p}{\mu}\frac{\partial p}{\partial x} = \frac{1}{2\mu}\frac{\partial p^2}{\partial x} = \frac{d}{dt}\left(\mu\frac{dx}{dt}\right).$$

Introducing Eq. (A-16) we obtain the equation of motion

$$\frac{d}{dt}(\mu v_x) = -\frac{\partial V}{\partial x} \tag{A-17}$$

and corresponding equations for the y and z coordinates. As these are the correct basic equations for the motion of a particle with variable mass, the principle of least action is verified for this case also.

Several Independent Variables.—The method of the calculus of variations can be applied when there are two or more ultimate independent variables. In this case we start with a multiple integral such as

$$J[u] = \int\int_G F(u, u_x, u_y, x, y)\, dx\, dy. \tag{A-18}$$

This integral is extended over a region G of the x,y plane on the boundary of which the undetermined function $u(x,y)$ is required to have predetermined values. u_x and u_y are the partial derivatives of u with respect to

x and y. Starting with a primary function $\varphi(x,y)$ which satisfies the boundary conditions we form the family of comparison functions

$$u(x,y,\alpha) = \varphi(x,y) + \alpha\eta(x,y)$$

as before. $\eta(x,y)$ is arbitrary except for continuity restrictions and requirement that it shall vanish on the boundary of G. Defining the first variation as before, we obtain

$$\delta J = \int\int_G\left[\frac{\partial F}{\partial u}\delta u + \frac{\partial F}{\partial u_x}\frac{\partial}{\partial x}(\delta u) + \frac{\partial F}{\partial u_y}\frac{\partial}{\partial y}(\delta u)\right]dx\,dy.$$

Integrating by parts this becomes

$$\delta J = \int\int_G\left[\frac{\partial F}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial u_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial u_y}\right)\right]\delta u\,dx\,dy$$
$$- \int_\Gamma\left[\frac{\partial F}{\partial u_x}\frac{dy}{ds} - \frac{\partial F}{\partial u_y}\frac{dx}{ds}\right]\delta u\,ds. \quad\text{(A-19)}$$

Γ is the boundary of G. Since δu vanishes on Γ, δJ vanishes, *i.e.*, J is stationary, when $\varphi(x,y)$ is a solution of the partial differential equation

$$\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial u_x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial u_y}\right) - \frac{\partial F}{\partial u} = 0. \quad\text{(A-20)}$$

Example II.—Consider the vibrations of a stretched string of elastic modulus λ and mass per unit length ρ. Let the string extend along the x axis from the origin to the point $x = l$. Let u denote the displacement in the direction of the y axis, the other components of displacement being assumed to vanish. Hamilton's principle requires that the integral

$$J[u] = \int_{t_0}^{t_1}(T - V)dt$$

shall be stationary for the natural motion. In this case

$$T = \int_0^l \tfrac{1}{2}\rho\dot{u}^2 dx,$$
$$V = \int_0^l \frac{1}{2}\lambda\left(\frac{\partial u}{\partial x}\right)^2 dx,$$

and we have

$$\delta\int_{t_0}^{t_1}dt\int_0^l dx[\tfrac{1}{2}\rho\dot{u}^2 - \tfrac{1}{2}\lambda u_x^2] = 0. \quad\text{(A-21)}$$

Equation (A-20) takes the familiar form

$$\rho\frac{\partial^2 u}{\partial t^2} = \lambda\frac{\partial^2 u}{\partial x^2}. \quad\text{(A-22)}$$

The calculus of variations and its applications to mechanics are discussed in a great many textbooks, but the writer is particularly indebted to the excellent treatment of the subject in Chap. IV of Courant and Hilbert's *Methoden der mathematischen Physik*, I Berlin, 1924 and 1931, and warmly recommends this discussion to the reader.

APPENDIX B

DERIVATION OF EQUATION (15·7)[1]

We desire to evaluate the function $\Psi(x,y,z,t)$ of Eq. (15·6) for values of x, y, z, t large enough to serve for momentum measurements. The components of the apparent, or most probable, momentum for any such set of values of x, y, z, t are

$$\xi = \frac{\mu x}{t}, \qquad \eta = \frac{\mu y}{t}, \qquad \zeta = \frac{\mu z}{t}.$$

It is convenient to introduce these along with t as the variables upon which Ψ depends. The probability of the momentum p_x, p_y, p_z will now be given by the limit of Ψ when t becomes infinite while ξ, η, ζ have the constant values p_x, p_y, p_z.

· To avoid cumbersome formulas we consider here only the one-dimensional case where Ψ is independent of y and z. There is no fundamental difficulty in extending the result to any number of dimensions. We suppose, then, that

$$\Psi(x,t) = \Psi\left(\frac{\xi t}{\mu},t\right) = \int_\infty G(\sigma)e^{2\pi i\left[\frac{\xi t\sigma}{\mu} - \frac{ht\sigma^2}{2\mu}\right]}d\sigma. \tag{B-1}$$

In order to throw the exponential into more convenient form we make the change of variables

$$u = \left(\frac{2ht}{\mu}\right)^{1/2}\left(\sigma - \frac{\xi}{h}\right). \qquad \alpha = \left(\frac{\mu}{2ht}\right)^{1/2}$$

Equation (B-1) now takes the form

$$\Psi\left(\frac{\xi}{2h\alpha^2},\frac{\mu}{2h\alpha^2}\right) = \alpha e^{\frac{\pi i\xi^2}{2h^2\alpha^2}}\int_{-\infty}^{+\infty} G\left(\frac{\xi}{h} + \alpha u\right)e^{-\frac{\pi i u^2}{2}}du. \tag{B-2}$$

It is convenient to split the integration into two parts covering the ranges $u > 0$ and $u < 0$, respectively. Introducing the Fresnel integrals[2]

$$C(u) = \int_0^u \cos\frac{\pi q^2}{2}\,dq - \frac{1}{2} = -\int_u^\infty \cos\frac{\pi q^2}{2}\,dq,$$

$$S(u) = \int_0^u \sin\frac{\pi q^2}{2}\,dq - \frac{1}{2} = -\int_u^\infty \sin\frac{\pi q^2}{2}\,dq,$$

[1] *Cf.* Sec. 15*b*, p. 62.
[2] The definitions of C and S here used differ by the additive constant $\frac{1}{2}$ from the customary definitions.

we have

$$\int_0^\infty G e^{-\frac{\pi i u^2}{2}} du = \int_0^\infty G \frac{d}{du}(C - iS) du.$$

Integrating by parts and observing that G vanishes when its argument becomes positively or negatively infinite, we obtain

$$\int_0^\infty G e^{-\frac{\pi i u^2}{2}} du = \int_0^\infty \frac{\partial}{\partial u}[G(C - iS)] du - \int_0^\infty (C - iS) \frac{\partial G}{\partial u} du$$

$$= \left(\frac{1 - i}{2}\right) G\left(\frac{\xi}{h}\right) - \alpha \int_0^\infty (C - iS) G'\left(\frac{\xi}{h} + \alpha u\right) du. \quad \text{(B-3)}$$

Here G' denotes the derivative of G with respect to its original argument. Similarly the range $u < 0$ gives

$$\int_{-\infty}^0 G e^{-\frac{\pi i u^2}{2}} du = \int_0^\infty G\left(\frac{\xi}{h} - \alpha u\right) e^{-\frac{\pi i u^2}{2}} du$$

$$= \left(\frac{1 - i}{2}\right) G\left(\frac{\xi}{h}\right) - \alpha \int_0^\infty (C - iS) G'\left(\frac{\xi}{h} - \alpha u\right) du. \quad \text{(B-4)}$$

Combining Eqs. (B-2), (B-3), and (B-4), we obtain

$$\alpha^{-1} e^{-\frac{\pi i \xi^2}{2 h^2 \alpha^2}} \Psi\left(\frac{\xi}{2 h \alpha^2}, \frac{\mu}{2 h \alpha^2}\right) = (1 - i) G\left(\frac{\xi}{h}\right) + \alpha \int_0^\infty (C - iS) \left[G'\left(\frac{\xi}{h} + \alpha u\right) \right.$$

$$\left. - G'\left(\frac{\xi}{h} - \alpha u\right) \right] du. \quad \text{(B-5)}$$

Let P denote the second term of the right-hand member of Eq. (B-5).

$$P(\xi, \alpha) = \alpha \int_0^\infty (C - iS) \left[G'\left(\frac{\xi}{h} + \alpha u\right) - G'\left(\frac{\xi}{h} - \alpha u\right) \right] du. \quad \text{(B-6)}$$

By introducing suitable hypotheses regarding the behavior of the function $\left[G'\left(\frac{\xi}{h} + \alpha u\right) - G'\left(\frac{\xi}{h} - \alpha u\right) \right]$ for large values of u it is possible to prove that $\lim_{\alpha \to 0} P(\xi, \alpha) = 0$. To this end we introduce αu as the variable of integration and denote it by the symbol v. Let the real and imaginary parts of $\left[G'\left(\frac{\xi}{h} + \alpha u\right) - G'\left(\frac{\xi}{h} - \alpha u\right) \right]$ be K_1 and K_2, respectively. Multiplying out the integrand, we can resolve P into the sum of four integrals of which the following is typical:

$$P_1 = \int_0^\infty K_1\left(\frac{\xi}{h}, v\right) C\left(\frac{v}{\alpha}\right) dv. \quad \text{(B-7)}$$

The function $C\left(\dfrac{v}{\alpha}\right)$ is of an oscillatory character with the asymptotic form

$$C\left(\frac{v}{\alpha}\right) = \frac{\alpha}{\pi v} \sin \frac{\pi v^2}{2\alpha^2}$$

for large values of the argument. $C(v/\alpha)$ approaches zero as α approaches zero for all values of v except zero, where it has the constant value $\frac{1}{2}$. At the same time the positive and negative arches get closer and closer together and by mutual cancellation tend to reduce the integral. Hence it is to be expected that if the integral $P_1(\xi,\alpha)$ exists at all [as it must if Eq. (B-1) is valid],

$$\lim_{\alpha \to 0} P_1(\xi,\alpha) = 0. \tag{B-8}$$

The author has not been able to devise a rigorous proof of the above statement for the general case but is indebted to the late Prof. O. D. Kellogg for the following proof for the special case that $K_1(\xi/h,v)$ is a function of *bounded variation*[1] in v. It follows from the method of deriving K_1 from $G(\sigma)$ that K_1 like G is analytic. Hence the requirement of a bounded variation reduces to the condition that $\displaystyle\int_0^\infty \left|\frac{\partial K_1}{\partial v}\right| dv$ shall exist. This restriction is very mild since $G(\sigma)$ is quadratically integrable. It permits us by a fundamental theorem to replace K_1 by the difference of two bounded, positive, monotonically decreasing functions which approach zero as v becomes infinite. Calling these functions $A(\xi,v)$ and $B(\xi,v)$, respectively, we have the relation

$$P_1(\xi,\alpha) = \lim_{b \to \infty} \left[\int_0^b C\left(\frac{v}{\alpha}\right) A(\xi,v) dv - \int_0^{b'} C\left(\frac{v}{\alpha}\right) B(\xi,v) dv \right]. \tag{B-9}$$

By the second theorem of the mean

$$\int_0^b C\left(\frac{v}{\alpha}\right) A(\xi,v) dv = A(\xi,a) \int_0^\theta C\left(\frac{v}{\alpha}\right) dv + A(\xi,b) \int_\theta^b C\left(\frac{v}{\alpha}\right) dv,$$

where $0 < \theta < b$. Due to the oscillatory character of $C(v/\alpha)$ the integral $\displaystyle\int_{v_1}^{v_2} C\left(\frac{v}{\alpha}\right) dv$ is the sum of an alternating series of terms representing the areas of the successive positive and negative arches of the integrand. Each of these terms is smaller than its predecessor and hence the sum is less than the area of the first complete arch to the right of the point $v = 0$. Let I denote the area of this largest arch:

$$I = \alpha \int_{0.406}^{3.28} C(z) dz,$$

[1] *Cf.* HOBSON, *Theory of Functions of a Real Variable,* edition of 1907, p. 256.

and let M denote the upper bound of $A(\xi,v)$. Then,

$$\left| A(\xi,a) \int_0^\theta C\left(\frac{v}{\alpha}\right) dv \right| \leqslant MI$$

and

$$\left| \int_0^b C\left(\frac{v}{\alpha}\right) A(\xi,v) dv \right| \leqslant 2MI.$$

Consequently, if b becomes infinite,

$$\left| \int_0^\infty C\left(\frac{v}{\alpha}\right) A(\xi,v) dv \right| \leqslant 2MI.$$

Next let α approach zero. Then I goes to zero, and hence

$$\lim_{\alpha \to 0} \int_0^\infty C\left(\frac{v}{\alpha}\right) A(\xi,v) dv = 0.$$

The same argument applies to the second term of Eq. (B-9), and so the theorem of Eq. (B-8) is proved.

As P_1 is one of four similar integrals into which the second term of Eq. (B-5) has been resolved, we infer that

$$\lim_{\alpha \to 0} \left[\alpha^{-1} e^{-\frac{\pi i \xi^2}{2h^2\alpha^2}} \Psi\left(\frac{\xi}{2h\alpha^2}, \frac{\mu}{2h\alpha^2}\right) \right] = (1-i) G\left(\frac{\xi}{h}\right).$$

Inserting the values of α and ξ we have that, for large values of t

$$\Psi(x,t) = \left(\frac{\mu}{2ht}\right)^{1/2} e^{\frac{\pi i \mu x^2}{ht}} (1-i) G\left(\frac{\mu x}{ht}\right).$$

This is a one-dimensional equivalent of Eq. (15·7) of Chap. II.

APPENDIX C

THEOREMS REGARDING THE LINEAR OSCILLATOR PROBLEM[1]

Several statements occur in the text of Sec. 19 which require a more rigorous discussion. The points in question are the following: (1) to show that $\psi(x)$ becomes infinite or approaches zero as x approaches either boundary; (2) to show that for a given α and E there is one, and only one, integral curve which vanishes at $x = -\infty$; (3), to show that solutions which vanish at $x = \pm\infty$ are quadratically integrable and hence satisfy the complete boundary condition A of Sec. 17. In this appendix we shall verify these statements and prove the integrability of $x^n|\psi|^2$, $\left|\dfrac{d\psi}{dx}\right|^2$, and $\left|\dfrac{d^2\psi}{dx^2}\right|^2$.

To prove point (1) we begin by observing that, since the curve is convex to the axis in H, its slope cannot change sign more than once. Consequently there exists a constant x_1 such that if $x > x_1 > x''$, $\psi(x)$ is monotone. Then either $\psi(x)$ becomes infinite with x or it approaches a finite limit, say B. We wish to prove that in the latter case $B = 0$. To this end we assume that $B \neq 0$ and seek a contradiction. Without loss of generality we may assume that B is positive. Then if $x > x_1$ the curve will be monotone decreasing and $\psi(x) > B$. By hypothesis, $V(x)$ is greater than E throughout the region H. Hence $\kappa(V - E)$ has a positive lower bound M in the region $x > x_1 > x''$. We conclude that in this region

$$\frac{d}{dx}\psi'(x) = \kappa(V - E)\psi > MB.$$

Then

$$\psi'(x) > \psi'(x_1) + MB(x - x_1).$$

Integrating again, we obtain

$$\psi(x) > \psi(x_1) + \psi'(x_1)(x - x_1) + \frac{MB}{2}(x - x_1)^2.$$

By hypothesis, $\psi(\infty)$ is finite. But this is clearly possible only if $B = 0$. Hence $\psi(x)$ must become infinite or approach zero as x approaches ∞. In exactly the same manner we can show that $\psi(x)$ becomes infinite, or approaches zero as x approaches $-\infty$.

[1] *Cf.* Sec. 19, p. 83.

568

For point (2) we identify α and β with the positive ordinate and the slope of the curve at some point $x = \xi$ in the region F, just as we did in the text. We consider the family of integral curves obtained by holding α and E constant, and varying β. No two of these curves can intersect to the left of ξ for, if they did, the difference between the two functions would be a solution of the equation for the same value of E which crosses the axis twice to the left of x'. Hence if $\beta' > \beta''$, we conclude that $\psi(\beta',x) < \psi(\beta'',x)$, to the left of ξ. For small positive values of β the curves will first approach the axis as x moves out along the negative axis, and then recede from it to become positively infinite as x approaches $-\infty$. For a sufficiently large positive value of β the curve is readily proved to cross the axis at some point x_1. As x_1 becomes increasingly negative, β must steadily decrease. Hence β approaches a limit β_1 as $x_1 \rightarrow -\infty$. Then the curve $\psi(\alpha,\beta_1,E,x)$ is a solution of the differential equation which satisfies the requirement that $\psi(-\infty) = 0$. For this curve cannot cross the x axis to the left of ξ because, if it did, β_1 could not be $\lim_{x_1 \rightarrow -\infty} \beta(x_1)$. Then it must approach the x axis as a limit. If it did not, we could find a curve coming still closer to the x axis without actually reaching it. This curve would have a larger β, and its existence would require that there be two curves for the same β, one crossing the axis the other not. Hence there is an integral curve for each pair of values of α and E which vanishes at $x = -\infty$. There can be only one, for otherwise the difference would be a solution of the differential equation crossing the axis in F but having a node at $-\infty$.

For the last point we proceed as follows. Consider the two functions $\psi_1(x)$ and $\psi_2(x)$ which have the same positive ordinate at $x = \xi$ and which vanish at $x = -\infty$. Let $\psi_1(x)$ satisfy the Eq. (18·2) and let $\psi_2(x)$ be a solution of

$$\psi_2'' + \kappa(E - V_2)\psi_2 = 0,$$

where V_2 is a constant greater than E but less than $V(x)$ when $x < \xi$. We assume that $\xi < x'$. Under these circumstances it follows that ψ_2 is greater than ψ_1 in the region $-\infty < x < \xi$. To prove the point we assume its converse

$$\psi_2 \leqslant \psi_1$$

to hold in a region M to the left of the point $x = \xi$. At the boundary points of M, say x_a and x_b, the two curves may be supposed to cross. As a special case, the region M may extend from $-\infty$ to ξ. From the differential equations satisfied by ψ_2 and ψ_1 we have

$$\frac{d}{dx}(\psi_1'\psi_2 - \psi_2'\psi_1) = \psi_1''\psi_2 - \psi_2''\psi_1 = \kappa(V - V_2)\psi_1\psi_2.$$

Hence

$$(\psi_1'\psi_2 - \psi_2'\psi_1)\Big|_{x_a}^{x_b} = \kappa \int_{x_a}^{x_b} (V - V_2)\psi_1\psi_2 dx.$$

The right-hand member of this equation is essentially positive and, since the curve ψ_2 cannot have the greater slope at x_a while ψ_1 cannot have the greater slope at x_b, the left-hand member cannot be positive. Hence we have a contradiction and no such region M can exist. We conclude that $\psi_2 > \psi_1$ at every point in the region $-\infty < x < \xi$.

It follows that if the integral

$$\int_{-\infty}^{\xi} \psi_2{}^2 dx$$

exists, the function ψ_1 must be quadratically integrable. But ψ_2 has the explicit form

$$\psi_2 = Ae^{\frac{2\pi x}{h}\sqrt{2\mu(V_2 - E)}}, \qquad x < \xi$$

from which it follows that the above integral does exist. A similar argument can be used to prove the quadratic integrability of any real discrete eigenfunction in the neighborhood of the right-hand boundary and hence over the complete fundamental interval $-\infty < x < +\infty$. Since the convergence of the integral is unaffected if we multiply ψ by any complex constant, we infer that, if ψ is any discrete eigenfunction, real or complex,

$$\int_{-\infty}^{+\infty} \psi\psi^* dx$$

exists.

It is equally clear that since $x^n\psi_1{}^2 < x^n\psi_2{}^2$ the convergence of $\int_{-\infty}^{\xi} x^n\psi_2{}^2 dx$ implies the convergence of $\int_{-\infty}^{\xi} x^n\psi_1{}^2 dx$. We readily infer that if ψ is any discrete eigenfunction, the product of ψ into any function of x which is continuous except at infinity, where it may have a pole or poles of finite order, is quadratically integrable. In particular, the integrals $\int_{-\infty}^{+\infty} V^2\psi\psi^* dx$ and $\int_{-\infty}^{+\infty} V\psi\psi^* dx$ are convergent.

If we multiply the differential equation (18·2) by ψ^* and integrate from $-\infty$ to $+\infty$ we obtain

$$\int_{-\infty}^{+\infty} \psi^* \frac{d^2\psi}{dx^2} dx = \kappa \left\{ \int_{-\infty}^{+\infty} V\psi\psi^* dx - E \int_{-\infty}^{+\infty} \psi\psi^* dx \right\}.$$

If ψ is a discrete eigenfunction with the eigenvalue E, the integrals on the right converge, and the integral on the left must converge also. But

$$\int_{-\infty}^{+\infty} \psi^* \frac{d^2\psi}{dx^2} dx + \int_{-\infty}^{+\infty} \frac{d\psi}{dx} \frac{d\psi^*}{dx} dx = \lim_{a \to \infty} \left(\psi^* \frac{d\psi}{dx} \right)_{x=a} - \lim_{b \to -\infty} \left(\psi^* \frac{d\psi}{dx} \right)_{x=b}.$$

Hence $d\psi/dx$ is quadratically integrable.

Finally we anticipate a result derived in Sec. 22 which states that if any two functions are quadratically integrable any linear combination of these functions will be quadratically integrable. Since it follows from the differential equation that $d^2\psi/dx^2$ is a linear combination of ψ and $V\psi$, we infer that $d^2\psi/dx^2$ is quadratically integrable.

APPENDIX D

MATHEMATICAL NOTES ON THE B. W. K. METHOD[1]

I. BEHAVIOR OF THE FUNCTIONS f_u AND f_v ON THE COMPLEX PLANE[2]

Consider first the special case of the parabolic potential curve

$$V(x) = \tfrac{1}{2}kx^2.$$

Introducing polar coordinates by means of the equations

$$x - x' = r'e^{i\theta'}, \qquad x - x'' = r''e^{i\theta''},$$

we express p^2 in the form

$$p^2 = -\mu k r' r'' e^{i(\theta' + \theta'')}.$$

As in footnote 2, p. 92, we assume a cut in the complex x plane along the axis of reals from x' to $+\infty$ and require that θ' and θ'' be restricted to the range $0 \leqslant \theta \leqslant 2\pi$. Let $p(x,E)$ and $p(x,E)^{-\frac{1}{2}}$ be defined by the equations

$$p = e^{-\frac{i\pi}{2}}(\mu k r' r'')^{\frac{1}{2}} e^{i\frac{(\theta' + \theta'')}{2}}, \qquad p^{-\frac{1}{2}} = e^{\frac{i\pi}{4}}(\mu k r' r'')^{-\frac{1}{4}} e^{-i\frac{(\theta' + \theta'')}{4}}. \tag{D-1}$$

The function $w(x,E)$ is defined by (21·6) with the provision that the path of integration shall not cross the cut. p, $p^{-\frac{1}{2}}$, and w are now uniquely defined at all points of the cut plane. The following expressions for these functions on different portions of the axis of reals are readily derived. (The regions F, G, H are shown graphically in Fig. 4, p. 81, and Fig. 6, p. 99.)

Upper lip of cut in H:

$$p = -i|p|, \qquad p^{-\frac{1}{2}} = e^{\frac{i\pi}{4}}|p^{-\frac{1}{2}}|, \qquad w = \frac{2\pi}{h}\int_{x'}^{x''}|p|d\xi - \frac{2\pi i}{h}\int_{x''}^{x}|p|d\xi.$$

Upper lip of cut in G:

$$p = |p|, \qquad p^{-\frac{1}{2}} = |p^{-\frac{1}{2}}|, \qquad w = \frac{2\pi}{h}\int_{x'}^{x}|p|d\xi.$$

In F:

$$p = i|p|, \qquad p^{-\frac{1}{2}} = e^{-\frac{i\pi}{4}}|p^{-\frac{1}{2}}|, \qquad w = -\frac{2\pi i}{h}\int_{x}^{x'}|p|d\xi.$$

[1] Cf. Sec. 21.
[2] Cf. p. 97.

Lower lip of cut in G:

$$p = -|p|, \qquad p^{-\frac{1}{2}} = -i|p^{-\frac{1}{2}}|, \qquad w = -\frac{2\pi}{h}\int_{x'}^x |p|d\xi.$$

Lower lip of cut in H:

$$p = -i|p|, \qquad p^{-\frac{1}{2}} = -e^{\frac{i\pi}{4}}|p^{-\frac{1}{2}}|, \qquad w = -\frac{2\pi}{h}\int_{x'}^{x''} |p|d\xi - \frac{2\pi i}{h}\int_{x''}^x |p|d\xi$$

The above restrictions on $p^{-\frac{1}{2}}$ and w afford unique definitions of $p^{-\frac{1}{2}}e^{iw}$ and $p^{-\frac{1}{2}}e^{-iw}$ at all points of the cut plane. For convenience we identify f_u and f_v with the portions of these functions spread out over the *upper* half of the cut plane. For the lower half of this plane we set

$$\tilde{f}_u = p^{-\frac{1}{2}}e^{iw}, \qquad \tilde{f}_v = p^{-\frac{1}{2}}e^{-iw}. \tag{D-2}$$

Then, $f_u, f_v, \tilde{f}_u, \tilde{f}_v$ are uniquely defined on the axis of reals, and for such real points we have the relations:

In F: $\qquad \tilde{f}_u = f_u, \qquad\qquad\qquad \tilde{f}_v = f_v. \tag{D-3}$

In G: $\qquad \tilde{f}_u = -if_v, \qquad\qquad\qquad \tilde{f}_v = -if_u. \tag{D-4}$

In H: $\qquad \tilde{f}_u = -f_u e^{-\frac{4\pi i}{h}\int_{x'}^{x''}|p|d\xi}, \qquad \tilde{f}_v = -f_v e^{\frac{4\pi i}{h}\int_{x'}^{x''}|p|d\xi}. \tag{D-5}$

f_u and $\tilde{f}_u$ both become infinite at each end of the axis of reals, while f_v and $\tilde{f}_v$ approach zero as x moves out to infinity along either the positive or negative axis of reals. For large values of $|x|$ the functions $f_u, \tilde{f}_u$ are *dominant*, $f_v, \tilde{f}_v$ are *subdominant*.

Further information can be derived from an examination of the lines of constant absolute value and of constant phase angle (argument) of the functions $e^{\pm iw}$. A line of each type can be drawn through every point of the complex x plane where p does not vanish. Separating x into real and imaginary parts and describing p by polar coordinates we write

$$x = u + iv, \qquad p = |p|e^{ix}.$$

The differential equation for the lines of constant absolute value is

$$\frac{dv}{du} = -\tan \chi = -\tan \frac{1}{2}(\theta' + \theta'' - \pi). \tag{D-6}$$

With the aid of this equation the general form of the lines is readily sketched in, as shown in Fig. 6, p. 99, Sec. 21d. p is real along the perpendicular bisector of the line $x'x''$, and hence $|e^{iw}|$ decreases as the point x moves away from the axis of reals on this line. At large distances from x' and x'' the exponential factor is the controlling one in determining the magnitude of the functions f_u and f_v. We conclude that the four cross-hatched regions in Fig. 6, have the following properties. In A and B f_u and $\tilde{f}_u$ are dominant while f_v, $\tilde{f}_v$ are subdominant. In C on the other hand, f_v is dominant while f_v is dominant in D.

As $\left|\dfrac{Q}{p^2}\right|$ and $\left|\dfrac{Q}{p}\right|$ become infinite at any simple zero of p^2, any curve on which f_u and f_v are good approximate solutions of (18·2) will have to keep away from the points x', x''. If such a curve is to connect the interval F of the axis of reals with the origin or with the interval H, it must inevitably penetrate one of the regions C and D.

In the more general case of an anharmonic oscillator, such as that contemplated in Fig. 5, the function $p^2/2\mu = E - V(x)$ will have pairs of conjugate complex roots in addition to the real roots x' and x''. These roots complicate the discussion somewhat, making it necessary to introduce new cuts radiating one from each complex zero to infinity in order to get single-valued functions f_u, f_v. In order to study the variations of a_u and a_v along a good path Γ by means of Eqs. (21·18) and (21·19) it is necessary to avoid discontinuities by choosing a path which does not cross any of these cuts. The behavior of the lines of constant $|e^{iw}|$ in the neighborhood of x' and x'' remains qualitatively the same, however, provided that all the other roots are relatively distant from the real pair. Hence any suitable path Γ for which $\mu_\Gamma \ll 1$ must enter C in passing around x' and x'' in the upper half plane.

II. EVALUATION OF THE LOW-ENERGY LEVELS OF A ONE-DIMENSIONAL OSCILLATOR BY THE B. W. K. METHOD[1]

In the case of the lower energy levels of a one-dimensional oscillator, one can establish the Bohr formula (21·10) by means of direct paths leading from the region A of Fig. 6 to B around x' and x'' without touching the axis of reals in G. Let us assume that the coefficients $a_u(x)$ and $a_v(x)$ are to be fitted to an eigenfunction $\psi_n(x, E_n)$ of Eq. (18·2) [cf. Eqs. (21·15) and (21·16)]. ψ_n and f_v both vanish at each end of the axis of reals, while f_u does not. Hence the coefficient $a_u(x)$ must vanish at $x = \pm \infty$ on the axis of reals.

We assume that the plane is cut in the manner suggested above, so that f_u, f_v are single-valued on the upper half-plane and $\bar{f}_u$, $\bar{f}_v$ on the lower half-plane. The values of a_u, a_v for the lower half-plane will be indicated by $\tilde{a}_u$ and $\tilde{a}_v$, respectively. On that portion of the axis of reals which we have labeled F ($x < x'$) the pairs of functions f_u, f_v and $\bar{f}_u, \bar{f}_v$ are equal. Hence the two sets of coefficients a_u, a_v and $\tilde{a}_u$, $\tilde{a}_v$ are equal on F for all ψ functions. On the other hand, no such equality holds *in general* on the cut portion of the axis of reals. The essential feature of our derivation is to show that if there exists a suitable good path Γ joining a pair of points P_1 in F and P_4 in H and crossing no cuts, the relations

$$\tilde{a}_u(x) = a_u(x) = 0, \qquad \tilde{a}_v(x) = a_v(x) \tag{D-7}$$

hold approximately for large real values of x in H, provided that $\psi = \psi_n$.

[1] *Cf.* p. 105.

We further assume that the inequality (21·32), or an equivalent inequality, holds for P_1, and that $|a_u(P_4)|$ is bounded in the same way. Under these circumstances we can neglect $a_u(P_1)$ and $a_u(P_4)$.

Since Γ is to cross no cuts, it will enclose no complex zeros of $E - V(x)$ between itself and the axis of reals. Let us choose for Γ a path lying wholly in the upper half-plane and denote its image in the axis of reals by $\tilde{\Gamma}$.

As f_v is dominant in the Stokes region C, a_v is sensibly constant over the portion of Γ which lies in C. Since a_u is negligibly small at P_1 and P_2, a_v will be sensibly constant on the portions of Γ in A and B. The reflected path $\tilde{\Gamma}$ has the same properties as Γ, and, as $\tilde{a}_v(P_1) = a_v(P_1)$, we infer that $\tilde{a}_v(P_4) = a_v(P_4)$ when small quantities of the order of μ_Γ are neglected. But,

$$a_v f_v(x) = \psi_n(x) = \tilde{a}_v \tilde{f}_v(x)$$

at $x = P_4$, and it follows from D-5 that

$$e^{\frac{4\pi i}{h} \int_{x'}^{x''} |p| d\xi \, \pm \, i\pi} = 1. \tag{D-8}$$

This leads at once to the quantum condition (21·10) for the eigenvalues of (18·2).

The accuracy of this result is limited only by the quality of the path Γ and the finite values of the constant N of (21·32) at P_1 and P_4. In the special case of an harmonic oscillator, where $V = \frac{1}{2}kx^2$, there are no complex zeros of $p(x,E)$ and the path can be pushed as far away from x' and x'' as desired. Since $\left|\dfrac{Q}{p}\right|$ and $\left|\dfrac{Q}{p^2}\right|$ vary as $|x|^{-3}$ and $|x|^{-4}$ for large values of $|x|$, it follows that by swinging around the origin in a wide enough arc we can reduce the error in $(D-8)$ below any assignable value. Thus the B. W. K. method yields a rigorous independent proof of the energy-level formula of the harmonic oscillator of Sec. 20.

III. TRANSMISSION OF MATTER WAVES AT A POTENTIAL BARRIER[1]

We assume that the potential function in the neighborhood of the "hill" is nearly parabolic in form, so that in first approximation in the neighborhood of the hill we can write

$$p^2 = 2\mu(E - V) = \mu k(x - x_1)(x - x_2) = \mu k r_1 r_2 e^{i(\theta_1 + \theta_2)},$$

where x_1, x_2 are the classical turning points of Fig. 7 and r_1, θ_1, r_2, θ_2 are the polar coordinates of $x - x_1$ and $x - x_2$, respectively. As in the "valley" case previously discussed, we cut the complex x plane along the axis of reals from $x = x_1$ to $x = +\infty$, restricting θ_1 and θ_2 to the

[1] *Cf.* p. 110.

range $0 \leqslant \theta \leqslant 2\pi$. Following the sign convention of p. 110 we write

$$p = |p|e^{\frac{i}{2}(\theta_1 + \theta_2)}, \quad p^{-\frac{1}{2}} = |p^{-\frac{1}{2}}|e^{-\frac{i}{4}(\theta_1 + \theta_2)}. \tag{D-9}$$

Let w be defined by Eq. (21·6) with the provision that the path of integration shall not cross the cut. As before we identify f_u and f_v with those portions of the functions $p^{-\frac{1}{2}}e^{iw}$ and $p^{-\frac{1}{2}}e^{-iw}$, respectively, which are spread out over the *upper* half of the cut plane. $\bar{f}_u, \bar{f}_v$ are the corresponding functions for the *lower* half of the cut plane.
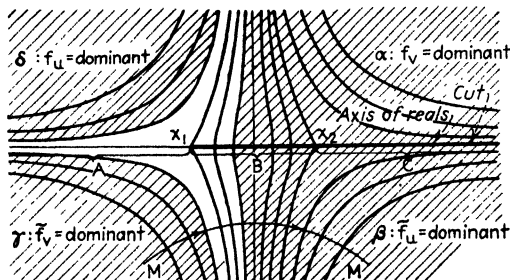
Fig. 27.—Level lines of the function $|e^{iw}|$ and Stokes regions for the problem of the parabolic potential hill.

If the turning points x_1, x_2 of the hill problem are made to coincide with the corresponding turning points x', x'' of the valley problem, the momentum p at every point of the cut plane in the hill problem is equal to the corresponding momentum for the valley problem multiplied by i. Thus the lines of constant $|e^{iw}|$ for the valley problem shown in Fig. 6, p. 99, become lines of constant argument in the hill problem. The new lines of constant absolute value are the orthogonals of the old set. They are shown in Fig. 27 together with one of the new lines of constant argument MM. The arrow heads on MM indicate the direction of increasing $|e^{iw}|$.

There are four Stokes regions α, β, γ, δ associated with the four successive quadrants. The corresponding dominant functions are indicated in the figure.

Let us assume that progressive matter waves are incident on the hill from the left side only. These will then be reflected as well as incident waves in the region A to the left of the hill in Fig. 7, but only emergent transmitted waves in the region C to the right of the hill. As noted on p. 110, the emergent waves on either side are of the f_u type. We therefore assume that a_v vanishes in the Stokes region α of the complex plane shown in Fig. 27, thereby insuring that it shall vanish on that portion of the axis of reals in the interval C where $\dfrac{h^2}{4\pi^2}\left|\dfrac{Q}{p^2}\right|$ is small. If there is a good path Γ extending from the interval A on the axis of reals to C,

passing around x_1 and x_2 in the upper half-plane, it will be a line on which a_u is sensibly constant.

It follows that there is a connection formula of the type

$$\underbrace{f_u + cf_v}_{A} \underset{C}{\overset{\leftharpoondown}{\leftarrow}} f_u \qquad \text{(D-10)}$$

in which the constant c is yet to be determined. By setting $\tilde{a}_v$ equal to zero in the Stokes region γ we obtain the companion formula

$$\underbrace{f_u = \tilde{f}_u}_{A} \rightarrow \underbrace{\tilde{f}_u + d\tilde{f}_v}_{C} \qquad \text{(D-11)}$$

for waves incident on the hill from the right. Both relations are equally good for high and low barriers. They have been derived on the assumption that the potential function is parabolic but are valid for hills of more general form, provided that the other roots of $E - V(x)$ are far enough away from x_1 and x_2 to permit the required good path Γ connecting the regions A and C on the axis of reals. In the region A:

$$\left.\begin{array}{l} f_u = \tilde{f}_u = |p|^{-\frac{1}{2}} \exp\left\{\dfrac{2\pi i}{h}\left[\displaystyle\int_x^{x_1}|p|d\xi + \dfrac{h}{4}\right]\right\}, \\[4mm] f_v = \tilde{f}_v = |p|^{-\frac{1}{2}} \exp\left\{-\dfrac{2\pi i}{h}\left[\displaystyle\int_x^{x_1}|p|d\xi - \dfrac{h}{4}\right]\right\}. \end{array}\right\} \qquad \text{(D-12)}$$

In the region C:

$$\left.\begin{array}{l} e^K f_u = e^{-K}\tilde{f}_u = |p|^{-\frac{1}{2}} \exp\left[\dfrac{2\pi i}{h}\displaystyle\int_{x_2}^x|p|d\xi\right], \\[4mm] e^{-K} f_v = e^K \tilde{f}_v = |p|^{-\frac{1}{2}} \exp\left[-\dfrac{2\pi i}{h}\displaystyle\int_{x_2}^x|p|d\xi\right], \end{array}\right\} \qquad \text{(D-13)}$$

where

$$K = \frac{2\pi}{h}\int_{x_1}^{x_2}|p|d\xi. \qquad \text{(D-14)}$$

It is useful to introduce at this point the equation of continuity

$$\operatorname{div}\vec{I} = -\mu\frac{\partial}{\partial t}(\Psi\Psi^*),$$

in which $\vec{I}$ is the mass current density defined for charged particles in three dimensions by Eq. (8·5) of Chap. I. In the case of monochromatic matter waves $\Psi\Psi^*$ is independent of the time and $\operatorname{div}\vec{I}$ must vanish. In a one-dimensional problem such as that under consideration, it follows that the current is constant for all real values of the basic coordinate x. Setting the vector potential of Eq. (8·5) equal to zero, we obtain the

expression

$$I[\psi] = \frac{h}{4\pi i}\left(\psi^*\frac{d\psi}{dx} - \psi\frac{d\psi^*}{dx}\right) \tag{D-15}$$

for the scalar value of the current in the direction of the positive x axis due to a single-energy wave function with space factor ψ. I is evidently merely a multiple of the Wronskian of the functions ψ and ψ^*.

It follows from (D-12) and (D-15) that in the region A on the axis of reals f_u and f_v represent currents of unit magnitude moving to left and right, respectively. Since $f_u = -f_v^*$, the currents are additive, *i.e.*, if α and β are any constants $I[\alpha f_u + \beta f_v] = I[\alpha f_u] + I[\beta f_v]$. In the region A the function ψ_a to which (D-10) refers can be regarded as the superposition of an incident current of magnitude $|c|^2$ and a reflected current of unit magnitude. In the region C the expressions $e^\kappa f_u$ and $e^{-\kappa}\bar{f}_u$ represent a unit current moving to the right, while $e^{-\kappa}f_v$ and $e^\kappa \bar{f}_v$ represent a unit current moving to the left. Equating the net currents on the two sides of the hill for the special case $\psi = \psi_a$, we obtain Eq. (21·40).

APPENDIX E

THE REDUCTION OF CERTAIN BOUNDARY-VALUE PROBLEMS BASED ON SELF-ADJOINT DIFFERENTIAL EQUATIONS TO VARIATIONAL FORM[1]

Variational Problem A.—The theorem a stated in the text (Sec. 24) can be proved as follows. Forming the first variation of the integral $J[y]$ defined by Eq. (24·2), we obtain

$$\delta J = \int_a^b [\delta y^*(\Lambda y + \lambda \rho y) + y^*(\Lambda \delta y + \lambda \rho \delta y)]dx. \qquad \text{(E-1)}$$

Since the operator Λ is self-adjoint, we can reduce this expression by means of Eq. (23·9), identifying the y of that equation with our present δy and the z with our present y^*. Using Eq. (24·5) we throw J into the form

$$\delta J = \int_a^b [\delta y^*(\Lambda y + \lambda \rho y) + \delta y(\Lambda^* y^* + \lambda \rho y^*)]dx$$

$$= \text{real part } \left\{ 2 \int_a^b \delta y^*(\Lambda y + \lambda \rho y)dx \right\}. \qquad \text{(E-2)}$$

This expression vanishes if y and λ form a solution of Eq. (24·1) so that the eigenfunctions and eigenvalues of Eq. (24·1) with the boundary conditions (α) are eigenfunctions and eigenvalues of problem A. Conversely, every solution of problem A is also a solution of Eq. (24·1), for if a function $y(x)$ and parameter-value λ do not form a solution of Eq. (24·1), we can always choose the arbitrary function δy^* so that $\delta J \neq 0$ and thus prove that y, λ do *not* form a solution of A.

It follows as a corollary that the stationary values of J are all zero.

Variational Problem B.—According to theorem b, p. 131, a necessary and sufficient condition that $\varphi(x)$ forms a solution of problem B is that $\varphi(x)$ and $\lambda_\varphi \equiv Q[\varphi]/N[\varphi]$ form together a solution of problem A.

It will first be proved that the condition is necessary. We adopt the notation introduced in Appendix A indicating that the variation of Q is based on the unvaried function $y(x)$ and the particular variation $\delta y = \delta \alpha \eta(x)$ by writing

$$\delta Q = \delta_y{}^\eta Q.$$

[1] *Cf. Sec.* 24, pp. 130–132.

Then, if $\varphi(x)$ is a solution of problem B,

$$\delta_{\varphi}\eta\left(\frac{Q}{N}\right) = \frac{1}{N[\varphi]}\left\{\delta_{\varphi}\eta Q - \frac{Q[\varphi]}{N[\varphi]}\delta_{\varphi}\eta N\right\} = 0 \qquad \text{(E-3)}$$

for every admissible $\eta(x)$. The corresponding variation of J is

$$\delta_{\varphi}\eta J = \delta_{\varphi}\eta\{-Q + \lambda N\} = -\delta_{\varphi}\eta Q + \lambda\delta_{\varphi}\eta N. \qquad \text{(E-4)}$$

If we give the adjustable constant λ the value

$$\lambda_{\varphi} = \frac{Q[\varphi]}{N[\varphi]}, \qquad \text{(E-5)}$$

it follows that $\delta_{\varphi}\eta J$ is zero for every η. Thus $\varphi(x)$, λ_{φ} constitute a solution of problem A.

Conversely, to prove the sufficiency of the condition, let us assume that the function $\zeta(x)$ is a solution of problem A with the eigenvalue λ_{ζ}. Since the stationary values of J are all zero,

$$J[\zeta] = -Q[\zeta] + \lambda_{\zeta}N[\zeta] = 0,$$

or

$$\lambda_{\zeta} = \frac{Q[\zeta]}{N[\zeta]}. \qquad \text{(E-6)}$$

Also

$$\delta_{\zeta}\eta J = -\delta_{\zeta}\eta Q + \lambda_{\zeta}\delta_{\zeta}\eta N = 0. \qquad \text{(E-7)}$$

Eliminating λ between (E-6) and (E-7) we obtain

$$\delta_{\zeta}\eta\left(\frac{Q}{N}\right) = -\frac{1}{N[\zeta]}\delta_{\zeta}\eta J = 0.$$

This completes the proof.

Variational Problem C.—Theorem c of Sec. 24 states that every solution φ of problems A and B, when normalized, yields a solution of problem C, and that conversely, every solution ζ of C is a solution of A and of B with the eigenvalue $\lambda_{\zeta} = -Q[\zeta]$ for the former problem.

Problem C asks for solutions of the variational equation $\delta Q = 0$ when the comparison functions $y(x)$ are subject to the normalization condition

$$N[y] \equiv \int_a^b \rho|y|^2 dx = 1. \qquad \text{(E-8)}$$

This type of problem lies outside the scope of the elementary theory of the calculus of variations given in Appendix A. It can be treated by the well-known method of undetermined multipliers due to Lagrange, or by the following direct attack.

According to the definition of Appendix A, δy is defined as

$$\left[\frac{\partial}{\partial\alpha}y(x,\alpha)\right]_{\alpha=0}\delta\alpha,$$

where $y(x,\alpha)$ is the one-parameter family of comparison functions

$$y(x,\alpha) = \varphi(x) + \alpha\eta(x). \tag{E-9}$$

δy is accordingly equal to $\eta(x)\delta\alpha$ and is arbitrary except for boundary and continuity conditions. The condition (E-8) applied to this function $y(x,\alpha)$ does not lead to a simple restriction on $\eta(x)$ and δy. We therefore replace (E-9) by the alternative family of comparison functions

$$\bar{y}(x,\alpha) = \frac{\varphi(x) + \alpha u(x)}{\sqrt{1 + \alpha^2 N[u]}}. \tag{E-10}$$

The normalization requirement is now seen to be satisfied if, and only if, $u(x)$ is orthogonal to $\rho\varphi$. A permissible first variation in y for problem C is now defined by

$$\delta_\varphi{}^u y = \left[\frac{\partial}{\partial\alpha}\bar{y}(x,\alpha)\right]_{\alpha=0}\delta\alpha = u(x)\delta\alpha, \tag{E-11}$$

where $u(x)$ is required to be orthogonal to $\rho\varphi$. With the aid of the restricted variations so defined we can proceed to the proof of theorem c.

It is clear that if $\psi(x)$ is a solution of either of the problems A and B, and if a is an arbitrary constant, $a\psi(x)$ is another solution with the same eigenvalue. It follows that every solution of the problems A and B can be normalized in accordance with (E-8). Let $\varphi(x)$ be such a normalized solution. Then $\delta_\varphi{}^u[Q/N] = 0$ for arbitrary functions $\eta(x)$ which satisfy the boundary and continuity conditions. If so, the restricted variation $\delta_\varphi{}^u[Q/N]$ calculated from (E-11) must also vanish, as $\delta_\varphi{}^u y$ is a special case of the unrestricted variation δy. Clearly,

$$\delta_\varphi{}^u N = \text{real part } (\rho\varphi, \delta_\varphi{}^u y) = 0. \tag{E-12}$$

Therefore

$$\delta_\varphi{}^u\left[\frac{Q}{N}\right] = \frac{1}{N[\varphi]}\left\{\delta_\varphi{}^u Q - \frac{Q[\varphi]}{N[\varphi]}\delta_\varphi{}^u N\right\} = \delta_\varphi{}^u Q = 0. \tag{E-13}$$

This proves that every normalized solution of problems A and B is also a solution of problem C.

We have next to verify the converse proposition that every solution of C is also a solution of A and B. Let $\zeta(x)$ be an arbitrary solution of C. Then the restricted variation $\delta_\zeta{}^u Q$ is zero for every $u(x)$ orthogonal to $\rho\zeta$. We have to show that the unrestricted variation $\delta_\zeta{}^\eta J$ vanishes for every $\eta(x)$ which satisfies the boundary-continuity conditions. With every unrestricted variation δy we associate the restricted variation

$$\delta_\zeta{}^u y = \delta y - \zeta(\rho\zeta, \delta y). \tag{E-14}$$

This function is actually orthogonal to $\rho\zeta$, for

$$(\rho\zeta, \delta_\zeta{}^u y) = (\rho\zeta, \delta y) - (\rho\zeta, \zeta)(\rho\zeta, \delta y) = 0.$$

Let us now form the unrestricted variation $\delta_\zeta{}^\eta J$. By Eqs. (E-2) and (E-13)

$$\delta_\zeta{}^\eta J = \text{real part } \left\{ 2\int_a^b [\delta_\zeta{}^u y + \zeta(\rho\zeta, \delta y)][\Lambda\zeta + \lambda\rho\zeta]dx \right\}.$$

Reducing the right-hand member, we obtain

$$\delta_\zeta{}^\eta J = -\delta_\zeta{}^u Q + \lambda\delta_\zeta{}^u N + \text{real part } \left(2(\rho\zeta, \delta y)\{-Q[\zeta] + \lambda N[\zeta]\} \right).$$

The first two terms drop out because $\zeta(x)$ is a solution of problem C. The expression in braces vanishes if we give λ the value

$$\frac{Q[\zeta]}{N[\zeta]} = Q[\zeta].$$

Thus $\delta_\zeta{}^\eta J$ vanishes for a proper choice of λ and $\zeta(x)$ is accordingly a solution of A. By theorem b it is also a solution of problem B. This completes the proof of the theorem.

THE LEGENDRE POLYNOMIALS AND ASSOCIATED LEGENDRE FUNCTIONS[1]

The definition and differential equation of the Legendre polynomials are given in Sec. 26. An equivalent definition is contained in the series expansion

$$\frac{1}{\sqrt{1 - 2x\tau + \tau^2}} = \sum_{n=0}^{\infty} P_n(x)\tau^n, \qquad \text{(F-1)}$$

which is of importance in potential theory.

The explicit formulas for the first five polynomials are

$$P_0(x) = 1, \qquad P_1(x) = x, \qquad P_2(x) = \frac{3}{2}x^2 - \frac{1}{2},$$

$$P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x, \qquad P_4(x) = \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8}.$$

The general formula is

$$P_n(x) = \frac{1 \cdot 3 \cdot 5 \cdots (2n - 1)}{n!}\left\{ x^n - \frac{n(n - 1)}{2(2n - 1)}x^{n-2} \right.$$
$$\left. + \frac{n(n - 1)(n - 2)(n - 3)}{2 \cdot 4(2n - 1)(2n - 3)}x^{n-4} - \cdots \right\} \qquad \text{(F-2)}$$

The last term in the braces is

$$+(-1)^{\frac{n}{2}}\frac{n!}{2 \cdot 4 \cdots n(2n - 1)(2n - 3) \cdots (n + 1)}$$

for even values of n and

$$+(-1)^{\frac{n-1}{2}}\frac{n! \, x}{2 \cdot 4 \cdots (n - 1)(2n - 1)(2n - 3) \cdots (n + 2)}$$

for odd values of n.

The more important properties of these polynomials are contained in the following formulas:

[1] *Cf.* Sec. 27, p. 145; Sec. 28, p. 150.

$$(n + 1)P_{n+1} = (2n + 1)xP_n - nP_{n-1}, \tag{F-3}$$

$$\frac{dP_n}{dx} \equiv P_n' = \frac{n(xP_n - P_{n-1})}{x^2 - 1}, \tag{F-4}$$

$$nP_n = xP_n' - P_{n-1}', \tag{F-5}$$

$$\int_{-1}^{+1} P_m(x)P_n(x)dx = \frac{2}{2n + 1}\delta_{mn}, \tag{F-6}$$

$$\int_{-1}^{+1} P_n(x)x^m dx = 0. \qquad m = 0, 1, 2, \cdots, n - 1 \tag{F-7}$$

The definition and differential equation of the associated Legendre functions are given in Sec. 28*d*. The associated Legendre functions of zero order are the ordinary Legendre polynomials just described. The first of the functions of higher order are tabulated herewith:

$$P_{1,1}(x) = (1 - x^2)^{1/2}, \qquad\qquad P_{2,2}(x) = 3(1 - x^2),$$
$$P_{2,1}(x) = 3x(1 - x^2)^{1/2}, \qquad\quad P_{3,2}(x) = 15x(1 - x^2),$$
$$P_{3,1}(x) = \tfrac{3}{2}(5x^2 - 1)(1 - x^2)^{1/2}, \qquad P_{3,3}(x) = 15(1 - x^2)^{3/2}.$$

The more important properties of these functions are given by Eqs. (F-8) to (F-15) below.

$$(l - \tau)(1 - x^2)^{1/2}P_{l,\tau} = xP_{l,\tau+1} - P_{l-1,\tau+1}, \tag{F-8}$$

$$(l + \tau + 1)(1 - x^2)^{1/2}P_{l,\tau} = P_{l+1,\tau+1} - xP_{l,\tau+1}, \tag{F-9}$$

$$(l + \tau)P_{l-1,\tau} = (2l + 1)xP_{l,\tau} - (l - \tau + 1)P_{l+1,\tau}, \tag{F-10}$$

$$(2l + 1)(1 - x^2)^{1/2}P_{l,\tau} = P_{l+1,\tau+1} - P_{l-1,\tau+1}, \tag{F-11}$$

$$(1 - x^2)^{1/2}P_{l,\tau} = 2(\tau - 1)xP_{l,\tau-1}$$
$$- (l + \tau - 1)(l - \tau + 2)(1 - x^2)^{1/2}P_{l,\tau-2}, \tag{F-12}$$

$$(1 - x^2)^{1/2}P_{l,\tau} = (l + \tau)xP_{l,\tau-1} - (l - \tau + 2)P_{l+1,\tau-1}, \tag{F-13}$$

$$(1 - x^2)\frac{d}{dx}P_{l,\tau} = (l + 1)xP_{l,\tau} - (l + 1 - \tau)P_{l+1,\tau}$$
$$= (l + \tau)P_{l-1,\tau} - lxP_{l,\tau} = -\tau xP_{l,\tau} + \sqrt{1 - x^2}P_{l,\tau+1}$$
$$= \tau xP_{l,\tau} - (l + 1 - \tau)(l + \tau)\sqrt{1 - x^2}P_{l,\tau-1}, \tag{F-14}$$

$$\int_{-1}^{+1} P_{l,\tau}P_{k,\tau}dx = \frac{2}{(2l + 1)}\frac{(l + \tau)!}{(l - \tau)!}\delta_{kl}. \tag{F-15}$$

The recurrence formulas are useful in deducing the functions of higher order and degree from the lower ones.

For further information regarding the associated Legendre functions the reader is referred to Bateman's *Partial Differential Equations of Mathematical Physics*, Chap. VI, Cambridge, 1932.

APPENDIX G

THE GENERALIZED LAGUERRE ORTHOGONAL FUNCTIONS[1]

The Laguerre polynomials are defined as the coefficients $L_k(x)$ in the expansion

$$\sum_{k=0}^{\infty} L_k(x)\frac{t^k}{k!} = \frac{e^{-\frac{xt}{1-t}}}{1-t}, \qquad t < 1 \tag{G-1}$$

or by the equivalent formula[2]

$$L_k(x) = e^x \frac{d^k}{dx^k}(x^k e^{-x}). \tag{G-2}$$

They obey the differential equation

$$xy'' + (1 - x)y' + ky = 0 \tag{G-3}$$

and have the following important properties:

$$L_{k+1}(x) - (2k + 1 - x)L_k(x) + k^2 L_{k-1}(x) = 0, \qquad k \geqslant 1 \tag{G-4}$$
$$xL_k'(x) = kL_k(x) - k^2 L_{k-1}(x), \qquad k \geqslant 1 \tag{G-5}$$
$$\int_0^{\infty} e^{-x}L_k(x)L_m(x)dx = (k!)^2 \delta_{km}. \tag{G-6}$$

The derivatives of the ordinary Laguerre polynomials[3] are called *generalized Laguerre polynomials* and are designated by the notation

$$L_k{}^n(x) = \frac{d^n}{dx^n}L_k(x). \tag{G-7}$$

The integer n is called the order of the polynomial. The generalized Laguerre polynomials satisfy the differential equation

$$xy'' + (n + 1 - x)y' + (k - n)y = 0, \tag{G-8}$$

and have the orthogonality property

$$\int_0^{\infty} e^{-x}x^n L_k{}^n(x)L_m{}^n(x)dx = \frac{(k!)^3}{(k - n)!}\delta_{km}. \tag{G-9}$$

Equation (G-9) is to be distinguished from the orthogonality relation

[1] *Cf.* Sec. 29c, p. 160.
[2] *Cf.* COURANT-HILBERT, *M.M.P.*, pp. 79, 284; RIEMANN-WEBER, *D.P.*, p. 341.
[3] *Cf.* COURANT-HILBERT, *M.M.P.*, p. 284; RIEMANN-WEBER, *D.P.*, p. 341; E. SCHRÖDINGER, *Ann. d. Physik* **80**, 483 (1926).

between $\mathfrak{R}_{nl}$ and $\mathfrak{R}_{n'l}$. One relation cannot be obtained from the other by direct transformation.

The first ten of these polynomials are

$$L_0{}^0(x) = 1, \qquad L_1{}^0(x) = -x + 1, \qquad L_1{}^1(x) = -1,$$
$$L_2{}^0(x) = x^2 - 4x + 2, \qquad L_2{}^1(x) = 2x - 4, \qquad L_2{}^2(x) = 2,$$
$$L_3{}^0(x) = -x^3 + 9x^2 - 18x + 6, \qquad L_3{}^1(x) = -3x^2 + 18x - 18,$$
$$L_3{}^2(x) = -6x + 18, \qquad L_3{}^3(x) = -6.$$

The explicit formula for the general case[1] is

$$L_k{}^n(x) = (-1)^k \frac{k!}{(k-n)!}\left\{ x^{k-n} - \frac{k(k-n)}{1!} x^{k-n-1} \right.$$
$$\left. + \frac{k(k-1)(k-n)(k-n-1)}{2!} x^{k-n-2} - \cdots \right\}. \quad \text{(G-10)}$$

The following integral formula derived by Schrödinger is also useful:

$$\int_0^\infty x^p e^{-x} L_k{}^n(x) L_{k'}{}^{n'}(x)\,dx = p!\,k!\,k'! \sum_{\tau=0}^{b} (-1)^{k+k'+\tau} \binom{p-n}{k-n-\tau}$$
$$\times \binom{p-n'}{k'-n'-\tau}\binom{-p-1}{\tau}. \quad \text{(G-11)}$$

Here b is the smaller of the two integers $k - n$ and $k' - n'$. The parenthesis symbols denote binomial coefficients and are defined by the relations

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \qquad n \geqslant k > 0$$
$$\binom{-n}{k} = (-1)^k \binom{n+k-1}{k},$$
$$\binom{n}{0} = \binom{n}{n} = 1. \qquad \binom{n}{k} = 0, \qquad k > n > 0.$$

In the special case that

$$n = n', \qquad k = k', \qquad p = n + 1,$$

all terms in the above sum vanish except those for which τ takes on the values $k - n$ and $k - n - 1$. Then (G-11) yields

$$\int_0^\infty x^{n+1} e^{-x} [L_k{}^n(x)]^2\,dx = (n+1)!\,(k!)^2 (-1)^{3k-n}\left\{ \binom{1}{0}^2 \binom{-n-2}{k-n} \right.$$
$$\left. - \binom{1}{1}^2 \binom{-n-2}{k-n-1} \right\} = \frac{(2k-n+1)(k!)^3}{(k-n)!}. \quad \text{(G-12)}$$

[1] *Cf.* Condon and Morse, *Q.M.*, p. 63.

Another special case of (G-11) of interest to us is that in which $p = n + 1$, $n' = n - 2$, $k' = k - 1$. Then,

$$\int_0^\infty x^{n+1} e^{-x} L_k{}^n(x) L_{k-1}{}^{n-2}(x) dx = (n + 1)! \, k! \, (k - 1)! \, (-1)^{k-n-1}$$

$$\times \left\{ \binom{1}{0}\binom{3}{1}\binom{-n - 2}{k - n} - \binom{1}{1}\binom{3}{2}\binom{-n - 2}{k - n - 1} \right\} =$$

$$- \frac{3(k!)^2 (k - 1)! \, (2k - n + 1)}{(k - n)!}. \quad \text{(G-13)}$$

The function $e^{-\frac{1}{2}x} x^{n/2} L_k{}^n(x)$ is called a generalized Laguerre orthogonal function of order n. It satisfies a Sturm-Liouville equation with eigenvalues ranging from n to infinity. Hence the Laguerre orthogonal functions of order n form a complete system. This is not true, however, of the radial wave functions $R_{nl}(r)$, as can be seen from the fact that the discrete eigenvalues of Eq. (29·1) have the upper limit zero. The explanation of this apparent discrepancy lies in the fact that the relation between the independent variables r and x involves the eigenvalue $n = \sqrt{-\epsilon}$ directly (*cf.* p. 160).

APPENDIX H

TWO THEOREMS RELATING TO THE CONTINUOUS SPECTRUM[1]

Theorem I.—*Let $y[x,\epsilon_n(b)]$ denote a real eigenfunction of the problem β of Sec. 30c; let $l_n(b)$ denote the distance from b to the preceding node of $y[x,\epsilon_n(b)]$; and let b' denote a suitably chosen point in the interval*

$$b - l_n(b) < x < b.$$

Then,

$$\Delta\epsilon(n,b) \equiv \epsilon_{n+1}(b) - \epsilon_n(b) = \frac{l_n(b)}{\kappa} \frac{y_x[b',\epsilon_n(b')]^2}{\int_a^{b'} y[x,\epsilon_n(b')]^2 dx} \cdot \quad \text{(H-1)}$$

Proof: It follows from the definitions of $l_n(b)$ and $y[x,\epsilon_n(b)]$ that

$$y[x,\epsilon_{n+1}(b)] = y[x,\epsilon_n(b - l_n(b))],$$

or that

$$\epsilon_{n+1}(b) = \epsilon_n(b - l_n(b)). \quad \text{(H-2)}$$

But by the first mean-value theorem

$$\epsilon_n(b) = \epsilon_n(b - l_n(b)) + l_n(b)\left(\frac{d\epsilon_n}{db}\right)_{b'}, \quad \text{(H-3)}$$

where b' lies in the interval $b - l_n(b) < x < b$. It follows that

$$\epsilon_{n+1}(b) - \epsilon_n(b) = -l_n(b)\left(\frac{d\epsilon_n}{db}\right)_{b'}. \quad \text{(H-4)}$$

$(d\epsilon_n/db)_{b'}$ can be evaluated by means of Green's formula

$$\left[y(x,E)y_x(x,E_1) - y(x,E_1)y_x(x,E) \right]_a^b = \kappa(E - E_1)\int_a^b y(x,E)y(x,E_1)dx. \quad \text{(H-5)}$$

We identify $y(x,E)$ with $y[x,\epsilon_n(b)]$, and $y(x,E_1)$ with a second member of a one-parameter family of integral curves which vary continuously with the second argument and satisfy the s.p.b.c. at $x = a$. Equation (H-5) becomes

$$y(b,E_1)y_x[b,\epsilon_n(b)] = \kappa[E_1 - \epsilon_n(a)]\int_a^b y[x,\epsilon_n(b)]y(x,E_1)dx. \quad \text{(H-6)}$$

Denoting the partial derivative of $y(x,E_1)$ with respect to the second

[1] *Cf.* Sec. 30d, p. 168.

argument by $y_E(x,E_1)$, we divide (H-6) by $[E_1 - \epsilon_n(b)]$ and allow E_1 to approach $\epsilon_n(b)$ as a limit. In this way we obtain

$$y_E[b,\epsilon_n(b)]y_x[b,\epsilon_n(b)] = \kappa \int_a^b y[x,\epsilon_n(b)]^2 dx. \qquad \text{(H-7)}$$

Differentiation of the identity

$$y[b,\epsilon_n(b)] \equiv 0$$

with respect to b yields

$$\frac{d\epsilon_n}{db} = -\frac{y_x[b,\epsilon_n(b)]}{y_E[b,\epsilon_n(b)]}. \qquad \text{(H-8)}$$

Combining (H-4), (H-7), and (H-8) we obtain

$$\epsilon_{n+1}(b) - \epsilon_n(b) = \frac{l_n(b)}{\kappa} \frac{y_x[b',\epsilon_n(b')]^2}{\int_a^{b'} y[x,\epsilon_n(b)]^2 dx}, \qquad \text{(H-9)}$$

as was to be proved.

Theorem II.—*Let $y(x,E)$ denote any family of integral curves of Eq. (30·10). For every positive value of E it is possible to choose constants A, x_0, γ, M, such that the B. W. K. approximation*

$$u(x,E) = A p^{-\frac{1}{2}} \cos \left\{ \frac{2\pi}{h} \int_{x_0}^x p(x',E)dx' + \gamma \right\}$$

conforms to the inequality

$$\left| \frac{y(x,E) - u(x,E)}{A} \right| < \frac{M}{x^2}. \qquad x > x_0 \qquad \text{(H-10)}$$

Proof: Having chosen a definite value of E we pick a value of x_0 greater than the largest root of $E - V(x)$. Let x_1 be an arbitrary point to the right of x_0 ($x_1 > x_0$). There is a B. W. K. approximation having the same ordinate and slope as $y(x,E)$ at x_1. Let

$$\tilde{u}(x,x_1) = \tilde{A}(x_1)p^{-\frac{1}{2}} \cos \left\{ \frac{2\pi}{h} \int_{x_0}^x p(x,E)dx + \tilde{\gamma}(x_1) \right\} \qquad \text{(H-11)}$$

denote this approximation. The difference

$$\xi(x,x_1) = y(x,E) - \tilde{u}(x,x_1) \qquad \text{(H-12)}$$

will then vanish with its first derivative with respect to x at x_1 and can readily be proved to satisfy the inhomogeneous differential equation

$$\frac{d^2\xi}{dx^2} + \kappa[E - V(x)]\xi = g(x,x_1,E), \qquad \text{(H-13)}$$

where

$$g(x,x_1,E) \equiv -\frac{\tilde{u}(x,x_1)}{4}\left[\frac{5}{4}\left(\frac{V'}{E-V}\right)^2 + \frac{V''}{(E-V)}\right] = -Q\tilde{u} \quad \text{(H-14)}$$

[*cf.* Eq. (21·12)].

Let $y_1(x,E)$, $y_2(x,E)$ denote any two real linearly independent solutions of (30·10). Their Wronskian, $W = y_1'y_2 - y_2'y_1$ is independent of x. The function

$$\varphi(x,x_1) = \frac{1}{W}\left[y_1\int_{x_1}^x gy_2 dx - y_2\int_{x_1}^x gy_1 dx\right] \quad \text{(H-15)}$$

is, like ξ, a solution of (H-13) which vanishes together with its first derivative at $x = x_1$. Hence ξ and φ are identical.

We desire to show that $\varphi(x,x_1)$ approaches a definite limit as x_1 becomes infinite. To this end we set up an upper bound for $|g(x,x_1,E)|$.

$|x^2 V'|$ and $|x^3 V''|$ have upper bounds D_1 and D_2, respectively, in the interval $x_0 < x < \infty$. Let ρ denote a positive real number less than unity. We *assume* that x_0 is so chosen that $E - V(x) > \rho E$ for all values of x greater than x_0. Then, if $x > x_0$,

$$|g(x,x_1,E)| < \frac{|\tilde{A}(x_1)|}{4(2\mu)^{1/4}(\rho E)^{5/4}}\left[D_2 + \frac{5}{4}\frac{D_1^2}{x\rho E}\right]\frac{1}{x^3}. \quad \text{(H-16)}$$

$\tilde{A}(x_1)$, $y_1(x,E)$, $y_2(x,E)$ are bounded, so that from (H-16) it follows that the integrals in (H-15) converge as $x_1 \to \infty$ and we see that

$$\lim_{x_1 \to \infty}[y(x,E) - \tilde{u}(x,x_1)] = y(x,E) - \lim_{x_1 \to \infty}\tilde{u}(x,x_1)$$

exists.

But if $\tilde{u}(x,x_1)$ approaches a definite limit as x_1 becomes infinite, $\tilde{A}$ and $\tilde{\gamma}$ must approach corresponding limits A and γ, respectively. We can accordingly identify the function $u(x,E)$, which appears in the statement of our theorem, with $\lim_{x_1 \to \infty}\tilde{u}(x,x_1)$.

By (H-15) and (H-16) we now have

$$\left|\frac{y(x,E) - u(x,E)}{A}\right| < \frac{1}{|W||A|}\left[|y_1|\int_x^\infty |g||y_2|dx + |y_2|\int_x^\infty |g||y_1|dx\right]$$

$$< \frac{N}{|W|\rho^{5/4}E^{5/4}}\left\{|y_1|\int_x^\infty \frac{|y_1|}{x^3}dx + |y_2|\int_x^\infty \frac{|y_2|}{x^3}dx\right\}, \quad \text{(H-17)}$$

with

$$N = \frac{1}{4(2\mu)^{1/4}}\left[D_2 + \frac{5D_1^2}{4x_0\rho E}\right]. \quad \text{(H-18)}$$

Since $|y_1|$ and $|y_2|$ are bounded in the region $x > 0$, Eq. (H-17) yields

$$\left|\frac{y(x,E) - u(x,E)}{A}\right| < \frac{M}{x^2}, \quad x > x_0 \quad \text{(H-19)}$$

where

$$M = \frac{N}{|W|(\rho E)^{9/4}} |y_1|_{\max} |y_2|_{\max}. \tag{H-20}$$

This proves the theorem.

It is to be observed that in general all the constants x_0, A, γ, M, which appear in the statement of the above theorem, vary with E. Only ρ can be chosen to be independent of E. It would be very useful if we could find an upper bound for M independent of E, but unfortunately such a bound does not exist. If we consider a case in which the potential energy approaches the value zero at infinity from below, we can treat x_0 as a constant. Then N varies as E^{-1} for small values of E. It follows from the B. W. K. approximation itself that $\dfrac{|y_1|_{\max}|y_2|_{\max}}{|W|}$ varies as $E^{-1/2}$. Hence, for small values of E, M varies as $E^{-\frac{11}{4}}$.

Corollary: If we differentiate (H-15) and then allow x_1 to become infinite, it is not difficult to prove that for any fixed energy

$$\left.\begin{aligned} \left|\frac{y'(x,E) - u'(x,E)}{A}\right| &< \frac{N_1}{x^2} \\ \left|\frac{y''(x,E) - u''(x,E)}{A}\right| &< \frac{N_2}{x^2} \end{aligned}\right\} \qquad x > x_0 \tag{H-21}$$

where N_1 and N_2 are positive constants.

CONCERNING THE EXPANSION OF Hf IN SPHERICAL HARMONICS[1]

We prove first the convergence of $\int_0^\infty r^2|G_{lm}|^2 dr$. Let R_{lm} be defined by the equation

$$f(r,\theta,\varphi) = G_{lm}Y_{lm} + R_{lm},$$

the notation being that of Sec. 30i. Since

$$G_{lm} = \iint fY_{lm}^* \sin \theta \, d\theta d\varphi,$$

we have $\iint R_{lm}Y_{lm}^* \sin \theta \, d\theta d\varphi = 0$. Hence

$$\iiint ff^* \sin \theta \, d\theta d\varphi = \iint |G_{lm}(r)|^2||Y_{lm}|^2 \sin \theta \, d\theta d\varphi + \iint |R_{lm}|^2 \sin \theta \, d\theta d\varphi$$

and

$$(f,f) = \int_0^\infty r^2 dr \int_0^\pi \sin \theta \, d\theta \int_0^{2\pi} ff^* d\varphi = \int_0^\infty r^2 \Big[|G_{lm}|^2 + \int_0^\pi \sin \theta \, d\theta \int_0^{2\pi} |R_{lm}|^2 d\varphi \Big] dr. \quad \text{(I-1)}$$

As f is quadratically integrable over x,y,z space, the integral $\int_0^\infty r^2|G_{lm}|^2 dr$ must converge.

The relation $F_{lm}(r) = r^{-1}\Lambda_l G_{lm}$, p. 176, will next be verified, thus establishing our right to apply H term by term to the series (30·37). By definition

$$F_{lm}(r) = \iint (Hf)Y_{lm}^* \sin \theta \, d\theta d\varphi = \iint Y_{lm}^*(HG_{lm}Y_{lm}) \sin \theta \, d\theta d\varphi$$
$$+ \iint Y_{lm}^*(HR_{lm}) \sin \theta \, d\theta d\varphi. \quad \text{(I-2)}$$

Let $(\mathfrak{L}^2)_{op}$ denote the angular-momentum operator

$$(\mathfrak{L}^2)_{op} = -\frac{h^2}{4\pi^2}\Big[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta}\Big(\sin \theta \frac{\partial}{\partial \theta}\Big) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \varphi^2} \Big]$$

of Eqs. (34·17) and (34·18). Then the operator H can be written in the form

$$H = -\frac{h^2}{8\pi^2\mu}r^{-2}\frac{\partial}{\partial r}\Big(r^2\frac{\partial}{\partial r}\Big) + V(r)\times + \frac{1}{2\mu r^2}(\mathfrak{L}^2)_{op}.$$

[1] *Cf.* Sec. 30i, p. 176.

Furthermore

$$(\mathfrak{L}^2)_{op} Y_{lm} = \frac{l(l+1)h^2}{4\pi^2} Y_{lm}$$

(*cf.* Sec. 28*d*), and hence

$$\int\int Y_{lm}{}^*(HG_{lm}Y_{lm}) \sin\theta\, d\theta d\varphi = -\frac{h^2}{8\pi^2\mu} r^{-2}\frac{\partial}{\partial r}\left(r^2\frac{\partial G_{lm}}{\partial r}\right) + V(r)G_{lm}$$
$$+ \frac{l(l+1)h^2}{8\pi^2\mu r^2}G_{lm} = r^{-1}\Lambda_l\mathcal{G}_{lm}. \quad \text{(I-3)}$$

The integral $\int\int Y_{lm}{}^*(HR_{lm}) \sin\theta\, d\theta d\varphi$ of (I-2) requires transformation. Thus

$$\int\int Y_{lm}{}^*(HR_{lm}) \sin\theta\, d\theta d\varphi = \frac{1}{2\mu r^2}\int\int Y_{lm}{}^*((\mathfrak{L}^2)_{op}R_{lm}) \sin\theta\, d\theta d\varphi$$

$$= \frac{1}{2\mu r^2}\int\int R_{lm}((\mathfrak{L}^2)_{op}Y_{lm}{}^*) \sin\theta\, d\theta d\varphi - \frac{1}{2\mu r^2}\int_0^{2\pi} d\varphi\int_0^{\pi} d\theta\left\{\frac{\partial}{\partial\theta}\left[\sin\theta\,(Y_{lm}{}^*\right.\right.$$
$$\left.\left.\times\frac{\partial R_{lm}}{\partial\theta} - R_{lm}\frac{\partial Y_{lm}{}^*}{\partial\theta}\right)\right] + \frac{\partial}{\partial\varphi}\left[\frac{1}{\sin\theta}\left(Y_{lm}{}^*\frac{\partial R_{lm}}{\partial\varphi} - R_{lm}\frac{\partial Y_{lm}{}^*}{\partial\varphi}\right)\right]\right\}.$$

Clearly

$$\int\int R_{lm}((\mathfrak{L}^2)_{op}Y_{lm}{}^*) \sin\theta\, d\theta d\varphi = \frac{l(l+1)h^2}{4\pi^2}\int\int R_{lm}Y_{lm}{}^* \sin\theta\, d\theta d\varphi = 0.$$

Since R_{lm} and $Y_{lm}{}^*$ are periodic in φ with period 2π,

$$\int_0^{2\pi} d\varphi\int_0^{\pi} d\theta\frac{\partial}{\partial\varphi}\left[\frac{1}{\sin\theta}\left(Y_{lm}{}^*\frac{\partial R_{lm}}{\partial\varphi} - R_{lm}\frac{\partial Y_{lm}{}^*}{\partial\varphi}\right)\right] = 0.$$

Finally

$$\int\int Y_{lm}{}^*(HR_{lm}) \sin\theta\, d\theta d\varphi = \frac{1}{2\mu r^2}\lim_{\theta_1\theta_2\to 0,\pi}\int_0^{2\pi}\left[\sin\theta\left(Y_{lm}{}^*\frac{\partial R_{lm}}{\partial\theta}\right.\right.$$
$$\left.\left.- R_{lm}\frac{\partial Y_{lm}{}^*}{\partial\theta}\right)\right]_{\theta=\theta_1}^{\theta=\theta_2}d\varphi = 0. \quad \text{(I-4)}$$

Combining (I-2), (I-3), and (I-4), we obtain the desired relation

$$F_{lm}(r) = r^{-1}\Lambda_l\mathcal{G}_{lm}.$$

The quadratic integrability of rF_{lm}, or $\Lambda_l\mathcal{G}_{lm}$ can now be established in the same manner as that of rG_{lm}.

APPENDIX J

THE JACOBI POLYNOMIALS[1]

The Jacobi polynomials $G(t)$ are terminating hypergeometric series and are related to the general hypergeometric functions

$$F(\alpha, \beta, \gamma, t) = 1 + \frac{\alpha\beta t}{1 \cdot \gamma} + \frac{\alpha(\alpha+1)\beta(\beta+1)t^2}{1 \cdot 2\gamma(\gamma+1)} + \cdots \qquad \text{(J-1)}$$

by the formula

$$G_p(a, b, t) = F(-p, a+p, b, t), \qquad \text{(J-2)}$$

where a, b, and p are integers. They are polynomials of degree p, as may be seen from the alternative definition

$$t^d(1-t)^s G_p(1+d+s, 1+d, t) = \frac{d!}{(d+p)!} \frac{d^p}{dt^p}[t^{d+p}(1-t)^{s+p}]. \qquad \text{(J-3)}$$

The polynomials $G_p(1+d+s, 1+d, t)$ are solutions of the differential equation

$$t(1-t)G'' + [1+d-t(d+s+2)]G' + p(p+d+s+1)G = 0. \qquad \text{(J-4)}$$

The normalization and orthogonality relation is

$$\int_0^1 t^d(1-t)^s G_p(1+d+s, 1+d, t)G_{p'}(1+d+s, 1+d, t)dt$$

$$= \delta_{pp'}\frac{p!(d!)^2(s+p)!}{(d+p)!(d+s+p)!(d+s+2p+1)}. \qquad \text{(J-5)}$$

The following relations are useful:

$$tG_p(1+d+s, 1+d, t) = \frac{d(d-1)}{(p+1)(d+s+p)}[G_{p+1}(d+s-1, d, t)$$
$$- G_{p+1}(d+s-1, d-1, t)], \qquad \text{(J-6)}$$

$$\frac{d}{dt}G_p(1+d+s, 1+d, t) = -\frac{p(1+d+s+p)}{1+d}G_{p-1}(3+d+s, 2+d, t), \qquad \text{(J-7)}$$

$$(1+d+s+2p)G_p(1+d+s, 1+d, t) = pG_{p-1}(2+d+s, 1+d, t)$$
$$+ (1+d+s+p)G_p(2+d+s, 1+d, t), \qquad \text{(J-8)}$$

$$(1+d+s+2p)(1+d)G_p(1+d+s, 1+d, t)$$
$$= -p(s+p)G_{p-1}(2+d+s, 2+d, t)$$
$$+ (1+d+s+p)(1+d+p)G_p(2+d+s, 2+d, t), \qquad \text{(J-9)}$$

$$G_p(1+d+s, 1+d, t) = \frac{d!(s+p)!}{s!(d+p)!}(-1)^p G_p(1+d+s, 1+s, 1-t). \qquad \text{(J-10)}$$

---

[1] *Cf.* Sec. 34*f*, p. 234.

The polynomials of lowest order are:

$$G_0(1+d+s, 1+d, t) = 1,$$

$$G_1(1+d+s, 1+d, t) = 1 - \frac{d+s+2t}{d+1}.$$

$$G_2(1+d+s, 1+d, t) = 1 - \frac{2(d+s+3)}{d+1}t + \frac{(d+s+3)(d+s+4)}{(d+1)(d+2)}t^2.$$

References

FRANK and VON MISES, *Differentialgleichungen der Physik*, I, 2d ed., p. 423, Berlin, 1930.

F. REICHE and H. RADEMACHER, *Zeits. f. Physik* **39**, 444 (1926), **41**, 453 (1927).

R. DE L. KRONIG and I. I. RABI, *Phys. Rev.* **29**, 262 (1927).

APPENDIX K

SCHLAPP'S METHOD[1]

The secular equation, $\det(^{(nm)}\mathbf{H}_1 - E^{(1)}\mathbf{I}) = 0$, can be written in the form

$$
\begin{vmatrix}
W & D_1 & 0 & . & . & . \\
D_1 & W & D_2 & . & . & . \\
0 & D_2 & W & . & . & . \\
. & . & . & . & . & . \\
. & . & . & . & W & D_{n-m-1} \\
. & . & . & . & D_{n-m-1} & W
\end{vmatrix} = 0, \tag{K-1}
$$

where $W = E^{(1)}/C$ and

$$
D_k{}^2 = \frac{k(2m + k)(n - m - k)(n + m + k)}{[2(m + k) - 1][2(m + k) + 1]}. \tag{K-2}
$$

In a determinant of the above type it is permissible to multiply any off-diagonal element by any number α not equal to zero, provided that at the same time we multiply the conjugate element by $1/\alpha$. Hence (K-1) is equivalent to

$$
\begin{vmatrix}
W & D_1{}^2 & 0 & . & . & . \\
1 & W & D_2{}^2 & . & . & . \\
0 & 1 & W & . & . & . \\
. & . & . & . & . & . \\
. & . & . & . & . & D_{n-m-1}{}^2 \\
. & . & . & . & 1 & W
\end{vmatrix} = 0. \tag{K-3}
$$

Transferring the first and last factors of the numerator of $D_k{}^2$, and the first factor of the denominator, to the conjugate element, we obtain a determinant of the form $K_{n,m}(W) \equiv$

$$
\begin{vmatrix}
W & \dfrac{(2m+1)(n-m-1)}{2m+1} & 0 & 0 & . \\[2ex]
\dfrac{1 \cdot (n+m+1)}{2m+3} & W & \dfrac{(2m+2)(n-m-2)}{2m+3} & 0 & . \\[2ex]
0 & \dfrac{2(n+m+2)}{2m+5} & W & \dfrac{(2m+3)(n-m-3)}{2m+5} & . \\[2ex]
0 & 0 & \dfrac{3(n+m+3)}{2m+7} & W & . \\[2ex]
. & . & & . & .
\end{vmatrix}
$$
$$
= 0. \tag{K-4}
$$

[1] *Cf.* Sec. 50, p. 407.

The sum of the two nonvanishing off-diagonal elements of the rth row is now $n - m - 1$. If we add to the elements of each column the sum of corresponding elements of every alternate succeeding column we obtain a determinant in which all elements lying below the principal diagonal have the values W, or $n - m - 1$, alternately. The elements on and above the principal diagonal are unaffected. Starting from the bottom we now subtract from each row the row next but one above it, thereby obtaining a new continuant in which the element (3,2) is zero. Thus $K_{n,m}(W) =$

$$
\begin{vmatrix}
W & n-m-1 & 0 & 0 & \cdot \\
n-m-1 & W & \dfrac{(2m+2)(n-m-2)}{2m+3} & 0 & \cdot \\
0 & 0 & W & \dfrac{(2m+3)(n-m-3)}{2m+5} & \cdot \\
0 & 0 & \dfrac{n+m+1}{2m+3} & W & \cdot \\
\cdot & \cdot & \cdot & \cdot &
\end{vmatrix} = 0.
$$

The new secular determinant factors into the product of the second-order minor

$$
\begin{vmatrix}
W & n-m-1 \\
n-m-1 & W
\end{vmatrix}
$$

and a determinant which is readily identified with $K_{n-1,m+1}(W)$. Thus $\pm(n - m - 1)$ are roots of $K_{n,m}(W)$. Similarly $\pm(n - m - 3)$ are roots of $K_{n-1,m+1}(W)$, etc. Proceeding in this way we finally obtain as the last factor of $K_{n,m}$ either W or $\begin{vmatrix} W & 1 \\ 1 & W \end{vmatrix}$ according as $n - m$ is even or odd. Hence the roots of $K_{n,m}(W)$ include every other integer from $-(n - m - 1)$ to $+(n - m - 1)$.

NAME INDEX

A

Abraham, 500
Auger, 213

B

Barnett, 501
Bartlett, 210
Bateman, 584
Bichowsky and Urey, 491
Bieberbach, 140, 142
Birkhoff, 25, 91, 105
Blaton (see Rubinowicz and Blaton)
Bôcher, 348
Bochner, 36
Bohr, 2, 72, 75, 90, 178, 179, 328, 374–378, 474, 479
Bolza, 133
Born, 2, 179, 379, 383, 427, 452, 482, 505
Born and Fock, 431
Born, Heisenberg, and Jordan, 370, 513, 543
Born and Jordan, 3, 348, 366, 371–373, 427
Born and Oppenheimer, 419
Breit, 232, 450, 525
Bridgman, 58, 76
Brillouin, L., 43, 91, 398, 477
Brillouin, M., 90
Brinkman, 465
de Broglie, 2, 3, 9, 10, 13, 19, 20, 90
Byerly, 149

C

Carleman, 252
Casimir, 179
Compton, 491
Condon (see Gurney and Condon)
Condon and Morse, 150, 222, 234, 451
Condon and Shortley, 338, 420, 462, 530, 535, 539, 540, 545, 556
Coolidge (see James and Coolidge)
Courant and Hilbert, 87, 90, 118, 122, 124, 130, 132, 135, 137–139, 145–147, 202, 216, 237, 410, 415, 563, 585

D

Dancoff and Inglis, 502
Darwin, 492, 510
Davisson and Germer, 14
Dennison, 535
Dirac, 79, 201, 241, 248, 265, 267, 281, 283, 287, 299, 315, 339, 427, 435, 441, 450, 452, 492, 509, 513, 545
Doi, 408
Drude, 8
Dunham, 107, 157, 388

E

Eckart, 308
(See also Hughes and Eckart)
Ehrenfest, 49
Einstein, 2, 4, 5, 76, 378, 448, 449
Einstein, Podolsky, and Rosen, 244, 328
Epstein, 403, 404, 408

F

Feenberg, 71, 98, 112
Fermi, 450, 477
Fischer, 259
Fock, 29, 312, 477
(See also Born and Fock)
Fowler, 439, 446, 448
Fowler and Nordheim, 109
Frank and von Mises, 595
Frenkel, 27, 34, 226, 376, 502, 525
Fues, 163, 177, 178, 195, 388
(See also Nordheim and Fues)
Furry, 244, 328

G

Gamow, 179, 192, 194
Gaunt, 525
Germer (see Davisson and Germer)
Gibbs, 55, 434
Gordon, 21

SUBJECT INDEX

A

Action function, 25, 44
Action integral, 8, 25, 560
Action variable, 375
Addition of states, 201
Adiabatic theorem, 431, 432
Adjoint manifold of operator, 276
Adjoint matrix, 350
Adjoint operator, first definition of, $122n$
   second definition of, 203, 512
Alkali spectra, energy levels of, 474–481
   theory of fine structure of, 503–507,
   519–522
Alpha particles, emission of, 179, 187–192
Analytic function, 140
Analyticity of wave functions, 18, 198, 200
Angular momenta, combination of, 495–
   498, 537–540
Angular momentum, 151, 224–234, 314–
   317
   of electron spin, 500, 503, 504, 510–519,
   529
   internal, 228
   resultant of spin and orbital, 498, 504,
   520
   of system of particles, 224, 227–234,
   292, 293
Angular-momentum operators, 224–234,
   292, 293, 314–317
Angular-momentum quantum number,
   151, 234, 493
Antisymmetric wave functions, 337, 531,
   533–536
Antisymmetrizer $\widehat{\mathsf{G}}$, 535, 536
Antisymmetrizing operator G, 338, 529,
   534, 535
Assemblages, canonical, 434, 446
   chaotic, 438
   concrete, 55, 319
   Gibbsian, 53–55, 329, 433–448
   mixed case, 54, 320
   pure case, 320
Asymptotic agreement of classical and
   quantum theories, 51, 229, 230, 302

Atomic energy levels, classification of,
   493, 494, 538, 539
Atomic model of Bohr, idealized, 474
Atomic units, 420
Atoms, complex, in problem A approxi-
   mation, 528
   in problem B approximation, 528,
   555
   in problem C approximation, 528
   with many electrons, 474–556
   perturbation theory of, 484–488,
   526–528
Auger effect, 195, 213, 214
Azimuthal quantum number l, 151

B

Bessel's inequality, 136
Binomial coefficients, 586
Bohr theory, 91, 108, 157, 178
   postulates of, 374
Boundary conditions, linear homo-
   geneous, 125, 130
   physical, 126, 153
   singular-point, 125–128, 130, 131, 153
Boundary and continuity conditions,
   78, 197–201
Brillouin-Wentzel-Kramers method, 46,
   90–112, 152, 155, 157, 158, 168, 182,
   572–578, 587
   higher approximations of, 107
   modification of, for radial motion,
   107, 108, 155
Broadened energy levels, 181, 183, 195

C

Calculus of variations, 130, 557–563
Canonical assemblage, 434, 446
Canonical equations (see Hamilton's
   canonical equations of motion)
Canonical Heisenberg matrices, 367
Canonical Schrödinger matrices, 369
Canonical transformation, 247, 355, 357,
   358

603